

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2004, 18:19:35 ; Search time 4623 seconds

(without alignments)
10247.445 Million cell updates/sec

Title: US-10-083-641a-6

Perfect score: 1093

Sequence: 1 gntatgnatcgttttaat.....ccaccaagttaggaccatnc 1093

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_ov:*
- 5: gb_ov:*
- 6: gb_ov:*
- 7: gb_ov:*
- 8: gb_ov:*
- 9: gb_ov:*
- 10: gb_ov:*
- 11: gb_ov:*
- 12: gb_ov:*
- 13: gb_ov:*
- 14: gb_ov:*
- 15: gb_ov:*
- 16: gb_ov:*
- 17: gb_ov:*
- 18: gb_ov:*
- 19: gb_ov:*
- 20: gb_ov:*
- 21: gb_ov:*
- 22: gb_ov:*
- 23: gb_ov:*
- 24: gb_ov:*
- 25: gb_ov:*
- 26: gb_ov:*
- 27: gb_ov:*
- 28: gb_ov:*
- 29: gb_ov:*
- 30: gb_ov:*
- 31: gb_ov:*
- 32: gb_ov:*
- 33: gb_ov:*
- 34: gb_ov:*
- 35: gb_ov:*
- 36: gb_ov:*
- 37: gb_ov:*
- 38: gb_ov:*
- 39: gb_ov:*
- 40: gb_ov:*
- 41: gb_ov:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	477	43.6	218817	2 AC114704	AC114704 Rattus no
2	477	43.6	240009	2 AC120292	AC120292 Rattus no
3	477	43.6	246796	2 AC094643	AC094643 Rattus no
4	373	34.1	1514	10 R069971	R069971 Rattus no
5	373	34.1	1560	10 BC062076	BC062076 Rattus no
6	119	10.9	152	10 AB010083	AB010083 Rattus no
7	44	4.0	1410	10 AB007143	AB007143 Mus muscu
8	44	4.0	1429	6 AR076190	AR076190 Sequence
9	44	4.0	1429	6 AR124103	AR124103 Sequence
10	44	4.0	1429	6 E23385	E23385 DNA encodin
11	44	4.0	247196	2 AC073822	AC073822 Mus muscu
12	26	2.4	51	6 AX165175	AX165175 Sequence
13	26	2.4	757	6 BD124885	BD124885 Primer fo
14	26	2.4	757	6 BD126291	BD126291 Primer fo
15	26	2.4	1087	9 BC039388	BC039388 Homo sapi
16	26	2.4	1365	12 AY335738	AY335738 Synthetic
17	26	2.4	1883	10 BC022165	BC022165 Mus muscu
18	26	2.4	2055	9 AB022341	AB022341 Homo sapi
19	26	2.4	2079	6 BD127287	BD127287 Primer fo
20	26	2.4	2079	6 AK074799	AK074799 Homo sapi
21	26	2.4	2105	6 BD176607	BD176607 Method of
22	26	2.4	2105	9 AB007144	AB007144 Homo sapi
23	26	2.4	2118	9 AK097643	AK097643 Homo sapi
24	26	2.4	2132	6 AR076189	AR076189 Sequence
25	26	2.4	2132	6 AR124102	AR124102 Sequence
26	26	2.4	2132	6 E23384	E23384 DNA encodin
27	26	2.4	2224	6 AX880020	AX880020 Sequence
28	26	2.4	2224	6 BD012208	BD012208 Novel gen
29	26	2.4	2224	6 BD158150	BD158150 Primer fo
30	26	2.4	2224	9 AK027590	AK027590 Homo sapi
31	26	2.4	2224	9 AC011488	AC011488 Homo sapi
32	26	2.4	162530	10 AC124439	AC124439 Mus muscu
33	26	2.4	170078	9 AC015914	AC015914 Homo sapi
34	26	2.4	194023	9 AC021541	AC021541 Homo sapi
35	26	2.4	210617	2 AC034201	AC034201 Homo sapi
36	26	2.4	224148	2 AC016179	AC016179 Homo sapi
37	26	2.4	228405	2 AC096406	AC096406 Rattus no
38	26	2.4	233704	2 AC103501	AC103501 Rattus no
39	25	2.3	184961	10 AC129021	AC129021 Mus muscu
40	25	2.3	195066	10 AC087900	AC087900 Mus muscu
41	24	2.2	1800	10 BC038491	BC038491 Mus muscu
42	24	2.2	1846	10 BC038491	BC038491 Mus muscu
43	24	2.2	2287	10 RAT66PK	RAT66PK Rat S6 prot
44	24	2.2	146759	10 AL604063	AL604063 Mouse DNA
45	24	2.2	243780	2 AC128095	AC128095 Rattus no

ALIGNMENTS

RESULT 1
AC114704/c 218817 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-55A15, WORKING DRAFT SEQUENCE.
DEFINITION AC114704
AC114704.4 GI:25136818
ACCESSION
VERSION
KEYWORDS HTG, HTGS PHASE2, HTGS DRAFT, HTGS FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 218817)
Muzny,D,Marle,,Metzker,M,lee,,Abramzon,S,,Adams,C,,Alder,J,,
Allen,C,,Allen,H,,Alsbrooks,S,,Amin,A,,Angiolano,D,,
Pred. No. is the number of results predicted by chance to have a

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F., Bismato, K., Blair, U., Blankenburg, K., Blyth, P., Brown, M., Bryan, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Chockley, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davis, M., Davis, C., Davy-Carroll, L., De Anda, C., Decker, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duva, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebreyes, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howell, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, U., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, D., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorens, L., Louised, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., McWhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Miosavljivic, A., Miner, G., Mirza, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokwelen, O., Okunolu, G., Olarnunpagon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Fu, L.-L., Piazio, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M.A., Reish, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.U., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soes, J., Steinle, M., Strong, R., Sutton, A., Svacek, A., Tabori, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yi, F., Zhang, U., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von Niedenhausen, A., Weiss, R., Smlth, D.R., Holt, R.A., Smlth, H.O., Weinstein, G., and Gibbs, R.A.

TITLE
Unpublished

JOURNAL
2 (bases 1 to 218817)

REFERENCE
Submitted (11-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

AUTHORS
3 (bases 1 to 218817)

TITLE
Rat Genome Sequencing Consortium.

JOURNAL
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
On Nov 20, 2002 this sequence version replaced gi:23265882. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSBS
Center clone name: CH230-55A15
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 207315 bases at least Q40
Consensus quality: 209091 bases at least Q30
Consensus quality: 210263 bases at least Q20
Estimated insert size: 213462; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 218817: contig of 218817 bp in length.
Location/Qualifiers
1. 218817
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-55A15"
/size="96483..97506"
/note="wgs contig"
misc_feature
217464..217908
/note="clone boundary
clone_end:sp6
site:
end_sequence:BH352650"

ORIGIN

Query Match 43.6%; Score 477; DB 2; Length 218817;
Best Local Similarity 100.0%; Pred. No. 6,9e-251;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

78 CCTCAGACGCGATTACCTCACTGATCTGTTGGTCCCGGTGCGGCGCAGCGT 137
Db CCTCAGACGCGATTACCTCACTGATCTGTTGGTCCCGGTGCGGCGCAGCGT 122471

138 CCTCTCCTCAAGGCAATCCCAAGTGTCTGTATGAGGCTCTTTGGGCACTTCTGTTGT 197
Db CCTCTCCTCAAGGCAATCCCAAGTGTCTGTATGAGGCTCTTTGGGCACTTCTGTTGT 122411

198 TGTGGAAACCTGGGAAACAGATGACAGAGGCTGGGTACAGAGTCTTGCTCTCTGG 257
Db TGTGGAAACCTGGGAAACAGATGACAGAGGCTGGGTACAGAGTCTTGCTCTCTGG 122351

258 GTTGCAGCGCTTGAAGTCTTCTTCCCAACAGCGGCGCAAGTGGCCATGTCGCAAGTG 317
Db GTTGCAGCGCTTGAAGTCTTCTTCCCAACAGCGGCGCAAGTGGCCATGTCGCAAGTG 122291

318 CCAGCAGAAAGGCAACCGGCAATGAGTACGGGCGCAAGTTATTAAGAAGGCGGCGCTGCC 377
Db CCAGCAGAAAGGCAACCGGCAATGAGTACGGGCGCAAGTTATTAAGAAGGCGGCGCTGCC 122231

378 GTTCAGCGCGCGCGGTGAGCCCTGAGAGATGAGCGGAGAGTGAAGCATCTTCGCGCA 437
Db GTTCAGCGCGCGCGGTGAGCCCTGAGAGATGAGCGGAGAGTGAAGCATCTTCGCGCA 122171

438 GATCGGCAACCCCAACATCATCAGCTGACAGATGTGTGAGAAACAAGACATGTGTGT 497

Db 122170 GATCGCGCACCCACATCATCATCGCTGCAGATGTTTCAGAGAACAGACGATGTGT 122111

Qy 498 GCTGATCTTGAGAGCTGCTCCGGCGGCGCAACTTTTCGACTTTTGGCTGAGAGAGA 554

Db 122110 GGTGATCTTGAGAGCTGCTCCGGCGGCGCAACTTTTCGACTTTTGGCTGAGAGAGA 122054

RESULT 2
AC120292 240009 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-239B15, WORKING DRAFT SEQUENCE, 2
DEFINITION Rattus norvegicus clone CH230-239B15, WORKING DRAFT SEQUENCE, 2
AC120292
AC120292
AC120292.4 GI:25008389
HTG: HTGS_PHASE1: HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 240009)
Muzny, D., Marie, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyadebe, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, J., Chen, G., Chen, R., Chen, Y., Chen, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, D., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dim, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, R., Jolivet, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C., L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loshushewa, L., Loulaged, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M., McNeill, T., Z., Meenen, E., Milsosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, N., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okunolu, G., Olariupusgon, A., Pal, S., Parks, K., Pascerik, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polinder, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C., D., Smaj, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, J., Valdes, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczek, R., Woodson, H., Wootley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, S., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE Direct Submission
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 240009)
AUTHORS Morley, K. C.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 240009)
REFERENCE Rat Genome Sequencing Consortium.
AUTHORS Direct Submission
TITLE Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 15, 2002 this sequence version replaced gi:23609710. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWXC
Center clone name: CH230-239B15
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 207207 bases at least Q40
Consensus quality: 208683 bases at least Q30
Estimated insert size: 213125; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOT: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOT: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 223747: contig of 223747 bp in length
* 223748 240009: gap of unknown length
* 223848 Location/Qualifiers
1. 240009
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-239B15"
complement (222684..223551)
/note="clone boundary
clone_end:17
site:
end_sequence:BZ106119"
complement (222705..223551)
/note="clone boundary
clone_end:17
site:
end_sequence:BZ106119"
complement (222705..223551)
/note="clone boundary
clone_end:17
site:
end_sequence:BZ106119"

ORIGIN
Query Match 43.6%; Score 477; DB 2; Length 240009;

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 244205: contig of 244205 bp in length
* 244206 244205: gap of unknown length
* 244306 245487: contig of 1182 bp in length
* 245488 245877: gap of unknown length
* 245878 246796: contig of 1209 bp in length.
Location/Qualifiers
1. 246796

FEATURES
source

misc_feature

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-5818"
10347. 199080
/note="clone boundary
clone end:Sp6
site:EcoRI
end_sequence:BH334507"
complement(241125..241693)
/note="clone boundary
clone end:T7
site:EcoRI
end_sequence:BH334504"

misc_feature

complement(241125..241693)
/note="clone boundary
clone end:T7
site:EcoRI
end_sequence:BH334504"

ORIGIN

Query Match 43.6%; Score 477; DB 2; Length 246796;
Best Local Similarity 100.0%; Pred. No. 6.8e-251;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 CCCTGACGACGGGATTAACCTGACTGTTCTTGGTCCCGGTCGGGCGCAGCGT 137
DB 235257 CCTCGACACGGGATTAACCTGACTGTTCTTGGTCCCGGTCGGGCGCAGCGT 235198
QY 138 CCTCTCCCTCAAGGCAATCCCAAGTGTCTGTCTGATGAGGCTCTTGGGCAAGTTCTGTTGT 197
DB 235197 CCTCTCCCTCAAGGCAATCCCAAGTGTCTGTCTGATGAGGCTCTTGGGCAAGTTCTGTTGT 235138
QY 198 TGTGGGAACCTGGGAAACAGATGCAAGAGGCTGGGGTCAAGAGTCTGCTCCCTG 257
DB 235137 TGTGGGAACCTGGGAAACAGATGCAAGAGGCTGGGGTCAAGAGTCTGCTCCCTG 235078
QY 258 GTCTGACAGGCTTAGTGTCTTCTCCCAAGCGGCGCAAGTTCGCATCGTGGCAAGTG 317
DB 235077 GTCTGACAGGCTTAGTGTCTTCTCCCAAGCGGCGCAAGTTCGCATCGTGGCAAGTG 235018
QY 318 CCAGCGAAGGCGACCGGCGATGAGATACCGGCGCAAGTTCTATTAAGAAGCGGCGCTGCG 377
DB 235017 CCAGCGAAGGCGACCGGCGATGAGATACCGGCGCAAGTTCTATTAAGAAGCGGCGCTGCG 234958
QY 378 GTCCAGCGGCGCGGTGTGAGCGGTGAGAGATGACGCGGAGTGAAGATCCTGCGCGA 437
DB 234957 GTCCAGCGGCGCGGTGTGAGCGGTGAGAGATGACGCGGAGTGAAGATCCTGCGCGA 234898
QY 438 GATCCGCGACCCCAACATCATCAAGCTGCAAGATGTGTTGAGAAACAAAGACAGATGTGT 497
DB 234897 GATCCGCGACCCCAACATCATCAAGCTGCAAGATGTGTTGAGAAACAAAGACAGATGTGT 234838
QY 498 GGTGATTTGAGAGTGTGTGCGGCGGCGAAGTTTTCAGTCTTGGTGAAGAGA 554
DB 234837 GGTGATTTGAGAGTGTGTGCGGCGGCGAAGTTTTCAGTCTTGGTGAAGAGA 234781

RESULT 4
RNO6971 1514 bp mRNA linear ROD 16-DEC-1998
LOCUS Rattus norvegicus mRNA for DAP-like kinase.
DEFINITION Rattus norvegicus mRNA for DAP-like kinase.
ACCESSION RNO6971.1 GI:3250894
VERSION RNO6971.1
KEYWORDS DAP-like kinase; dlk gene.

SOURCE
ORGANISM
Rattus norvegicus (Norway rat)

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
Kogel,D., Plotner,O., Landsberg,G., Christian,S. and
Scheidtmann,K.
TITLE
Cloning and characterization of Dlk, a novel serine/threonine
kinase that is tightly associated with chromatin and phosphorylates
core histones
JOURNAL
Oncogene 20, 2645-2654 (1998)

REFERENCE
AUTHORS
Kogel,D.
TITLE
Direct Submission
JOURNAL
Submitted (15-JUN-1998) Kogel D., Universitat Bonn, Institut fur
Genetik, Romerstr. 164, 53117 Bonn, GERMANY
REMARK
3 (bases 1 to 1514)
revised by [3]
REFERENCE
AUTHORS
Kogel,D.
TITLE
Direct Submission
JOURNAL
Submitted (22-JUN-1998) Kogel D., Universitat Bonn, Institut fur
Genetik, Romerstr. 164, 53117 Bonn, GERMANY

FEATURES
source
1. 1514
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/cell_line="SV52"
/cell_type="rat fibroblasts transformed with SV40 large T
antigen"

gene
CDS

1. 1514
/gene="dlk"
114. 1460
/gene="dlk"
/codon_start=1
/product="DAP-like kinase"
/protein_id="CAA07360.1"
/db_xref="GI:3250895"
/db_xref="GOA:O88764"
/db_xref="SPTREMBL:O88764"
translation="MSTFRQEDVDHYEMGEELSSGQFAIVKCOQKGTGEMVAKFI
KRRIPSSRGVSEIEEREYSILREHNIILHDPVFNKTDVVLILEVSGELF
DLAEKSLIDEDATFOILQIDGVHLSRIKHPDKENIMLTKHASPRIKLI
DPAKHSITDEASERKNI FGTPAPAPITVAVPEIGLADWMSIGVITLISGASPL
GETROETLINSAVNYDFDEYFSTSELAKDFIRLLVDDPKRMTIASSIEMSWIK
VRRREDARKERRRLPARLREYSLSHSSMPENISYASFRSRYLEDVAAEQGL
RELORGRQCEERYCALRVAAEQREARCRGSGAGLGRDLRLRLTELTGRTALRRAGE
EAAALLAGAGLKRRLCRLERNYDALAAQVAAEVQFVVDIVRALEORLAEQVR"

ORIGIN

Query Match 34.1%; Score 373; DB 10; Length 1514;
Best Local Similarity 99.8%; Pred. No. 2e-193;
Matches 493; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 288 CAGCGCGAGTTCGCGCATCGTCCGCAAGTGCACAGAAAGGCGACCGGCATGAGTAGCG 347
DB 173 CAGCGCGAGTTCGCGCATCGTCCGCAAGTGCACAGAAAGGCGACCGGCATGAGTAGCG 232
QY 348 GGCCAGTTCATTAAGAAAGCGCGCGCTCCGTCAGCGCGGCGGCTGAGACCGTGAGGA 407
DB 223 GGCCAGTTCATTAAGAAAGCGCGCGCTCCGTCAGCGCGGCGGCTGAGACCGTGAGGA 292
QY 408 GATCAGGCGGAGGTGAGCATCTGCGGAGATCCGCCAACCACCAACATCATCAGCTGCA 467
DB 293 GATCAGGCGGAGGTGAGCATCTGCGGAGATCCGCCAACCACCAACATCATCAGCTGCA 352
QY 468 CGATGTGTTGAGAAACAAGACAGATGTGTGCTGATCTTGAAGTGTGTCTCCGCGGCGGA 527
DB 353 CGATGTGTTGAGAAACAAGACAGATGTGTGCTGATCTTGAAGTGTGTCTCCGCGGCGGA 412
QY 528 ACTTTGCACTTTGGGCTGAGAAAGA-TCACTGACAGAGGATGAGCGCAGCGATTCTCT 586

Db 413 ACTTTGACTTCTTGCTGAGAGAGATGATGACAGAGATGAGGCCACGCACTTCT 472
 QY 587 CAAGCAGATCTGAGCGGTGCTCACTGACATCCAGCGCATGGCGCATTTGACCT 646
 Db 473 CAAGCAGATCTGAGCGGTGCTCACTGACATCCAGCGCATGGCGCATTTGACCT 532
 QY 647 GAAGCCGAGAAATCATGTTGTTGCAAGCATGACGCCAGCCAGCATTTAGCTCAT 706
 Db 533 GAAGCCGAGAAATCATGTTGTTGCAAGCATGACGCCAGCCAGCATTTAGCTCAT 592
 QY 707 CGACTTGGATCGCGCAGATCGAGCGCGGTGAGGATTCAGAAATCTTTGGAC 766
 Db 593 CGACTTGGATCGCGCAGATCGAGCGCGGTGAGGATTCAGAAATCTTTGGAC 652
 QY 767 GCCAGATTGCTGCG 780
 Db 653 GCCAGATTGCTGCG 666

RESULT 5
 BC062076 1560 bp mRNA linear ROD 12-NOV-2003
 DEFINITION Rattus norvegicus Death-associated like kinase, mRNA (cDNA clone
 MGC:72441 IMAGE:5599435), complete cds.
 ACCESSION BC062076
 VERSION BC062076.1 GI:38304023
 KEYWORDS MGC.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1560)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Statchenko, M., Soares, M.B., Bonaldo, A.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,
 Adnan, R.D., Muliyil, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shewchenko, Y.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shewchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butler, Y.S., Krzywinski, M.I., Skalska, J., Smalins, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL MEDLINE 22389257
 PUBMED 12477932
 2 (bases 1 to 1560)
 Strausberg, R.
 Title Direct Submission
 Submitted (10-NOV-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Jeff Green/Patruu Kondalah, NCI.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Gaithersburg, Maryland/
 Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov
 Akter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, B., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskett, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAC Plate: 137 Row: n Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 11968141.
 Location/Qualifiers
 1..1560
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="MGC:72441 IMAGE:5599435"
 /tissue_type="Prostate, pool of NCI_CGAP_P29, 40, 41 and
 NCI CGAP_P29, 39, 42"
 /clone_1lb="NCI CGAP_P50"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6.1"
 1..1560
 /gene="Dapkl"
 /db_xref="LocusID:64391"
 /db_xref="RCD:621766"
 115..1461
 /codon_start=1
 /product="Death-associated like kinase"
 /protein_id="AA062076.1"
 /db_xref="GI:38304024"
 /db_xref="LocusID:64391"
 /translation="MSFRQEDVEDYEMGEELSGQFAIVRCKQKGTGMEYAFI
 KKRRLPSRSGRESEIREVSLIREPHNITLHDVFNKTDVLLILEVSGEELF
 DPLAEKSLFEDAEATFOIKOILGVYHLSKRIATHDIDKENTMLDKAASPRITLI
 DPGIARHIEAGSEFKNIEGPEPVABEIVYVEPLGIEADWMSICVYTIILSSASPL
 GKEQETLITISVNTDFEYFSTSELANKDIRLVKDPKRRKTIKQSLHSWIK
 VRRREDARPERRRLRAAEIRVLSKSHSMPNCTSVASFERSVLEVAAGELG
 REAQRQRQCRERCAIRVAEQRERCRSGAGLGRDRLRTEIGRETAETRAQE
 EAPAAALGAGGLRRLRLRNRYDALAAGAAVGVPRDLVRLAEQERLQAEQVR"
 148..939
 /note="S.TKc: Region: Serine/Threonine protein kinases,
 catalytic domain. Phosphotransferases of the serine or
 threonine-specific kinase subfamily. The enzymatic
 activity of these protein kinases is controlled by
 phosphorylation of specific residues in the activation
 segment of the catalytic domain, sometimes combined with
 reversible conformational changes in the C-terminal
 autoregulatory tail"
 /db_xref="CDD:c000180"

ORIGIN
 Query Match 34.1%; Score 373; DB 10; Length 1560;
 Best Local Similarity 99.8%; Pred. No. 2e-193;
 Matches 493; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 288 CAGCGCCAGTGGCCCATCTGTCGCAAGTCCAGCAAGAGGACCGCGCATGAGTACCG 347
 Db 174 CAGCGCCAGTGGCCCATCTGTCGCAAGTCCAGCAAGAGGACCGCGCATGAGTACCG 233
 QY 348 GGCAGATTATTAAGAAGAGCGGCGCTGCGCTCACCGCGCGGTGAGCGATGAGA 407
 Db 234 GGCAGATTATTAAGAAGAGCGGCGCTGCGCTCACCGCGCGGTGAGCGATGAGA 293
 QY 408 GATCGAGCGCGAGGTGAGCATCTCTGCGCGAGATCCGCCACCCCAATCATCATCAGCTTGA 467
 Db 294 GATCGAGCGCGAGGTGAGCATCTCTGCGCGAGATCCGCCACCCCAATCATCATCAGCTTGA 353
 QY 468 CGATGTTCTGAGAAACAAGACAGATGGTGGTCTGATCTTGAGACTGTGTGTCGGCGGCGGA 527

Db 354 CAGTGTTCGAGAACACAGATGTGTGCTGATCTTGAGAGCTGTGTGTCCGGGGGGA 413
QY 528 ACTTTTCGACTTTTCTGCTGAGAGAGA-TCACTGACAGAGATGAGGCCAGGCACTTCT 586
Db 414 ACTTTTCGACTTTTCTGCTGAGAGAGAGTCACTGACAGAGATGAGGCCAGGCACTTCT 473
QY 587 CAGACGATCTTGGAGAGGTGTCCACTGACTGACCTGAGGCCATCGGCACTTTGACT 646
Db 474 CAGACGATCTTGGAGAGGTGTCCACTGACTGACCTGAGGCCATCGGCACTTTGACT 533
QY 647 GAGCCGAGAACATCATGTGTCTGACACAGATGAGGCCAGGCCAGGCACTTTAGCTCAT 706
Db 534 GAGCCGAGAACATCATGTGTCTGACACAGATGAGGCCAGGCCAGGCACTTTAGCTCAT 593
QY 707 CGACTTTGGCATGCGCCACAGATGAGGCCAGGCCAGGCACTTTAGCTCAT 766
Db 594 CGACTTTGGCATGCGCCACAGATGAGGCCAGGCCAGGCACTTTAGCTCAT 653
QY 767 GCAGAGTTCGTG 780
Db 654 GCAGAGTTCGTG 667
RESULT 6
AB010083 152 bp mRNA linear ROD 12-JUN-1998
LOCUS Rattus norvegicus mRNA for PKA alpha, partial cds.
DEFINITION AB010083
ACCESSION AB010083.1 GI:3219221
VERSION PKA alpha; kinase domain of PKA alpha.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 152)
AUTHORS Nara,K.
TITLE A novel protein kinase, PKA alpha, expressed in postnatal 7-day rat hippocampus
JOURNAL Published Only in DataBase (1998)
REFERENCE 2 (bases 1 to 152)
AUTHORS Nara,K.
TITLE Direct Submission
JOURNAL Submitted (29-DEC-1997) Kiyomitsu Nara, Mitsubishi Kasei Institute of Life Sciences, Department of Glycobiology, Minamioya 11, Machida, Tokyo 194, Japan (E-mail:nara@libra.1s.m-kagaku.co.jp, Tel:81-427-24-6298, Fax:81-427-24-6316)
FEATURES
Source Location/Qualifiers
1..152
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wister ST"
/db_xref="taxon:10116"
/tissue_type="hippocampus"
CDS
1..152
/function="putative serin/threonine protein kinase"
/note="kinase domain of PKA alpha"
/codon_start=1
/product="PKA alpha"
/protein_id="BAA28810.1"
/db_xref="GI:3219222"
/translation="HCDFKPENIMLDKHAASPRIKLIDFGIAHRIEAGSEFNINIGT
PAYFAP"
ORIGIN
Query Match 10.0%; Score 119; DB 10; Length 152;
Best Local Similarity 100.0%; Pred. No. 2e-53;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 648 AAGCGGAGAACATCATGTGTCTGACACAGATGAGGCCAGGCCAGGCACTTTAGCTCATC 707
Db 13 AAGCGGAGAACATCATGTGTCTGACACAGATGAGGCCAGGCCAGGCACTTTAGCTCATC 72

QY 708 GACTTTGGCATGCGCCACAGATGAGGCCAGGCCAGGCACTTTAGCTCATC 766
Db 73 GACTTTGGCATGCGCCACAGATGAGGCCAGGCCAGGCACTTTAGCTCATC 131
RESULT 7
AB007143 1410 bp mRNA linear ROD 25-FEB-1998
LOCUS Mus musculus mRNA for ZIP-kinase, complete cds.
DEFINITION AB007143
ACCESSION AB007143.1 GI:2911153
VERSION ZIP-kinase.
KEYWORDS ZIP-kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Kawai,T., Matsumoto,M., Takeda,K., Sanjo,H. and Akita,S.
TITLE ZIP kinase, a novel serine/threonine kinase which mediates apoptosis
JOURNAL Mol. Cell. Biol. 18 (3), 1642-1651 (1998)
MEDLINE 94847805
PUBMED 94848481
REFERENCE 2 (bases 1 to 1410)
AUTHORS Akita,S. and Kawai,T.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1997) Shizuo Akita, Hyogo College of Medicine, Department of Biochemistry, Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan (E-mail:akita@hyo-med.ac.jp, Tel:+81-798-45-6357, Fax:+81-798-46-3164)
FEATURES
Source Location/Qualifiers
1..1410
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
CDS
10..1356
/codon_start=1
/product="ZIP-kinase"
/protein_id="BAA2954.1"
/db_xref="GI:2911154"
/translation="MSTFROEDVEDHYEMGEELSGQFAIVKCCQKGTMEYAKRT
KKRLPSRGRVSRERIEEYSLREIRHPNITLHDVFENKDVVILIELVSGELF
DLAEKSLTEDEATOFKOTLDGVYHLKRIAPHDLKPEINIMLDKHAASPRIKL
DGIYAHRIEAGSEFNIGTPEFVAPRIVAVPEPLADMSIGVITYILSGSPF,
GETQOETITNISAVNYPDDEYFSTSEIAKDFIRLLVMDPKRMTIAGLSHWIK
VRREDGAKRKRRRLRLARKLRYSLSKSHSVPKNTYASFSRSRVLEDAVAAPQGL
REIQRGRQCRERVCALRAAEQREARCRDSAGLGRLLRLTELTGTEALRTBAQZ
EAFALILGAGGLKRRRLCRLENRYDALAAVAEVOFVADLVRALEQRLQAEQVR"
polYA_site
1410
/note="19 a nucleotides"
ORIGIN
Query Match 4.0%; Score 44; DB 10; Length 1410;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 415 CCGGAGTGAGCATCTGCGGAGATCGGCCAGGCCACCAATCAT 458
Db 196 CCGGAGTGAGCATCTGCGGAGATCGGCCAGGCCACCAATCAT 239
RESULT 8
AR076190 1429 bp DNA linear PAT 30-AUG-2000
LOCUS Sequence 4 from patent US 5958748.
DEFINITION AR076190
ACCESSION AR076190
VERSION AR076190.1 GI:10002936
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1429)

AUTHORS Akira,S. and Kawai,T.
TITLE DNA coding for serine/threonine kinase
JOURNAL Patent: US 5958748-A 4 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..1429
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 4.0%; Score 44; DB 6; Length 1429;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
196 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 239

RESULT 9
AR124103 1429 bp DNA linear PAT 16-MAY-2001
LOCUS AR124103
DEFINITION Sequence 4 from patent US 6171841.
ACCESSION AR124103
VERSION AR124103.1 GI:14109464
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1429)
AUTHORS Akira,S. and Kawai,T.
TITLE DNA coding for serine/threonine kinase
JOURNAL Patent: US 6171841-A 4 09-JUN-2001;
FEATURES Location/Qualifiers
source 1..1429
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 4.0%; Score 44; DB 6; Length 1429;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
196 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 239

RESULT 10
E23385 1429 bp DNA linear PAT 18-JUN-2001
LOCUS E23385
DEFINITION DNA encoding serine/threonine kinase.
ACCESSION E23385
VERSION E23385.1 GI:13024387
KEYWORDS JP 1999098984-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1429)
AUTHORS Shizuo,S. and Taro,K.
TITLE DNA encoding serine/threonine kinase
JOURNAL Patent: JP 1999098984-A 2 13-APR-1999;
COMMENT SCIENCE & TECH AGENCY
OS unidentified
PN JP 1999098984-A/2
PD 13-APR-1999
PF 26-SEP-1997 JP 1997261589

PI SHIZUO SHITARA TARO KAWAI
PC C12N15/09,C12N1/21,C12N9/12//C12N15/09,C12R1:91), (C12N1/21,
PC C12R1:19),
PC (C12N9/12,C12R1:19), C12N15/00, (C12N15/00,C12R1:91) CC
Strandedness: Double;
CC Topology: Linear;

FEATURES
source 1..1429
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 4.0%; Score 44; DB 6; Length 1429;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
196 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 239

RESULT 11

AC073822 247196 bp DNA linear HTG 18-JUL-2000
LOCUS AC073822
DEFINITION Mus musculus clone RP23-85K13, WORKING DRAFT SEQUENCE, 16 ordered
ACCESSION AC073822
VERSION AC073822.2 GI:9256816
KEYWORDS HTG, HTGS_PHASE2; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 247196)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Sequencing of Mouse
JOURNAL Unpublished
AUTHORS DOE Joint Genome Institute.
DIRECT SUBMISSION 2 (bases 1 to 247196)
JOURNAL DOE Joint Genome Institute.
COMMENT Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced GI:8810439.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1769765
Center clone name: RPO1-23_85K13

Summary Statistics
Consensus quality: 236727 bases at least Q40
Consensus quality: 243596 bases at least Q30
Consensus quality: 244906 bases at least Q20
Estimated insert size: 243000; agarose-1p estimation
Estimated insert size: 246496; sum-of-contigs estimation
Quality coverage: 7.76 in Q20 bases; agarose-1p estimation
Quality coverage: 7.65 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 18947: contig of 18947 bp in length
* 18948 19047: gap of unknown length
* 19048 20569: contig of 1522 bp in length
* 20570 27716: gap of unknown length
* 27717 27816: contig of 7047 bp in length
* 27817 29329: gap of unknown length
* 29330: contig of 1513 bp in length

*	29330	28429:	gap of unknown length
*	29430	62474:	contig of 33045 bp in length
*	62475	62574:	gap of unknown length
*	62575	66039:	contig of 3465 bp in length
*	66040	66139:	gap of unknown length
*	66140	91228:	contig of 25089 bp in length
*	91229	91328:	gap of unknown length
*	91339	11560:	contig of 24132 bp in length
*	11561	11560:	gap of unknown length
*	11561	13267:	contig of 17007 bp in length
*	13268	13267:	gap of unknown length
*	13268	136189:	contig of 3522 bp in length
*	136190	136289:	gap of unknown length
*	136290	178101:	contig of 48182 bp in length
*	178102	178201:	gap of unknown length
*	178202	187292:	contig of 9991 bp in length
*	187293	187392:	gap of unknown length
*	187393	188720:	contig of 1328 bp in length
*	188721	188820:	gap of unknown length
*	188821	193055:	contig of 4335 bp in length
*	193056	193155:	gap of unknown length
*	193156	205418:	contig of 12263 bp in length
*	205419	205518:	gap of unknown length
*	205519	247196:	contig of 41678 bp in length

ORIGIN

Query Match	4.0%	Score 44	DB 2	Length 247196
Best Local Similarity	100.0%	Pred. No.	2	3e-12
Matches 44	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	415	CGCGAGGTGAGCATCTCTGCGCGAGATCGGCCACCCCAATCAT	458
Db	154652	CGCGAGGTGAGCATCTCTGCGCGAGATCGGCCACCCCAATCAT	154695

RESULT 12	AX165175/c	LOCUS	AX:65175	51 bp	DNA	PAT 22-JUN-2001
DEFINITION	Sequence 370 from Patent WO0128586.				linear	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

variation

/note="single nucleotide polymorphism
Accession number cg43919086"

Query Match	2.4%	Score 26	DB 6	Length 51
Similarity	100.0%	Pred. No.	0.038	
Best Local	0	Mismatches	0	Gaps 0
Matches 26	Conservative	0	Indels	

```

QY      911 ATGTGAGCATCGGCGTCATCACCCTA 936
          |||||
Db      40  ATGTGAGCATCGGCGTCATCACCCTA 15

```

RESULT 13				
BD124885				
LOCUS				
DEFINITION	BD124885	757 bp	DNA	linear
	Primer for synthesizing full-length cDNA and use thereof.			
				PAT 18-SEP-2002

REFERENCE

1 (bases 1 to 757)
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Otsuka, E., Nakagaki, T., Wada, Y., Matsumoto, T., and Otsuki, T.

PC
PI TETSUJI OTSUKI, HISASHI KOGA
SHINTCHI KOJIMA,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
ISHII,
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
FUKUSHI, TETSUO UENO, TETSUO UENO, TETSUO UENO,
PI

```

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
10', C12P21/02, C12Q1/68//C12P21/09, G06F17/30 C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT source 1..757
/organism='Homo sapiens (human)'.

```

```

source
1..757
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

Query Match      2.4%;   Score 26;   DB 6;   Length 757;
Best Local Similarity 100.0%;   Pred. No. 0.031;
Matches      26;   Conservative      0;   Mismatches      0;   Indels      0;   Gaps      0

```

QY	579	CAGTTCCTCAGCAGATCCTGGACGG	604
Db	421	CAGTTCCCTCAAGCAGATCCTGGACGG	446

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
BD126291	BD126291	BD126291	BD126291.1	UP 2002017375-A/1722.	Homo sapiens (human)
	Primer for synthesizing full-length cDNA and use thereof.				
	757 bp	DNA	linear		
					PAT 18-SEP-2002

REFERENCE AUTHORS	TITLE
1 (bases 1 to 757) Oka, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.	Primer for synthesizing full-length cDNA and use thereof Patent: JP 2002017375-A 1722 22-JAN-2002;

HELIIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/1722
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO ISHII
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOTAJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12P21/68//C12P21/09, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..757
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.4%; Score 26; DB 6; Length 757;
Best Local Similarity 100.0%; Pred. No. 0.031; 0; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGACATCTCGACGG 604
|||||
421 CAGTTCTCAAGACATCTCGACGG 446
|||||

RESULT 15
BC039388/c
LOCUS BC039388 1087 bp mRNA linear PRI 06-NOV-2002
DEFINITION Homo sapiens, clone IMAGE:5298774, mRNA.
ACCESSION BC039388
VERSION BC039388.1 GI:24658265
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1087)
Strausberg, R.
Direct Submission
Submitted (01-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Miklos Paloytcs, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdgspxl1.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 77 Row: C Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Excellent library of
origin.
Location/Qualifiers

FEATURES
source

source 1..1087
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="5298774"
/tissue_type="testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN
Query Match 2.4%; Score 26; DB 9; Length 1087;
Best Local Similarity 100.0%; Pred. No. 0.03; 0; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 923 GCGTCATCACTACATCTGTGAGT 948
|||||
275 GCGTCATCACTACATCTGTGAGT 250
|||||

RESULT 16
AY335738
LOCUS AY335738 1365 bp mRNA linear SYN 15-OCT-2003
DEFINITION Synthetic construct Homo sapiens death-associated protein kinase 3
(DAPK3) mRNA, partial cds.
ACCESSION AY335738
VERSION AY335738.1 GI:33304114
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
1 (bases 1 to 1365)
REFERENCE
Park, J., Rolfs, A., Hu, Y., Shen, B., Vanberg, F., Morita, D.,
Keller, T., Zhu, D., Raphael, J., Bagui, M., Jepson, D., Harlow, E.,
Labaer, J., and Brizuela, L.
Cloning of human full-length CDS FLEXGene kinases in
recombinational vector system
Unpublished
2 (bases 1 to 1365)
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
AUTHORS
COMMENT
This CDS clone is a part of a collection of human full-length
expression clones generated by Harvard Institute of Proteomics.
Each CDS has been cloned without stop-codon (to allow fusion with
C-terminal tag). The CDS has been directionally cloned using BD
In-Fusion (TM) cloning system between the SalI and HindIII sites of
the pDNR-Dual vector. Additional sequences in the clone: 'ACC'
after SalI site and before 'ATG' to provide Kozak consensus
sequence; 'GG' after last codon and before HindIII site to maintain
reading frame.
Location/Qualifiers
1..1365
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone_image="5298774"
/tissue_type="placenta and brain"
/clone_lib="First strand cDNA from placenta and brain"
/lab_host="DH5alpha TI resistant"
/note="Vector: pDNR-Dual"
1..1365
/gene="DAPK3"
1..1365
/gene="DAPK3"
/note="Mutations: Stop->Leu"
/codon_start=1
/transl_table=11
/product="death-associated protein kinase 3"

gene
CDS

ORIGIN

Query Match 2.4%; Score 26; DB 12; Length 1365;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAGACAGATCCTGACG 604
 DB 352 CAGTTCCTCAGACAGATCCTGACG 377

/protein_id="AA02565.1"
 /db_xref="GI:33304115"
 /translation="MSTRQEDVDHVMKEBLGGGPAIVKRCQKGTGEYAKRI
 KRRRSSRRGVSRREIEREVNITLHDIPEKRTDVLLLEVSGLP
 DFLAKESITDEDAEPLKQILDGVHLHSHKRAHFDKPNIMLDKNVNPRIKLI
 DGIAMKIEAGNEFNIPTGPEVAPETVAYEPLGLADMSIGVITILLGASPL
 GETTQETLNTSAVYDDEEYFNTSELAQDFRLVQPKPKMTIAQSLSEHWK
 AIRRNVDGSGCRKPERRLKTPKTKYTLKSHSLPPNNSYADPFERSSVLEAA
 AEEGRLRECRSRICHDVEYALAAIYEKEMVRESLSQDLRLAQELTKTEAK
 RQAEPAKGLITSGLRKFRSLNRYEALAKQVASEKRFQDLVRALEQKQGYE
 CGLRL"

RESULT 17 BC022165 1883 bp mRNA linear ROD 08-OCT-2003
 LOCUS Mus musculus death-associated kinase 2, mRNA (cDNA clone MGC:13742
 DEFINITION IMAGE:4016911), complete cds.

ACCESSION BC022165
 VERSION BC022165.1 GI:18381096
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1883)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scherz, T.E., Brownstein, M.J., Usdin, T.B., Tashyuk, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullen, S.J., Bosak, S.A., McKean, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y.K., S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Rodigues, S.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Wyers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE JOURNAL MEDLINE
 PUBMED 22388257
 2 (bases 1 to 1883)
 AUTHOR Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (25-JAN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: gcaps-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>

FEATURES

source
 contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAP Plate: 18 Row: f Column: 7.
 Location/Qualifiers

gene

CDS

1..1883
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CEBCH 11"
 /db_xref="taxon:10090"
 /clone="MGC:13742 IMAGE:4016911"
 /tissue type="Mammary tumor metastasized to lung. Tumor
 arose spontaneously from a senescent normal mammary
 (clonal) outgrowth infected with the virus MMV."
 /clone_1db="NCI CGAP_Lu29"
 /lab_host="DHI0B"
 /note="vector: pCMV-SPORT6"
 1..1883
 /gene="Dabk2"
 /db_xref="locusid:13143"
 /db_xref="MGI:1341297"
 91..1203
 /codon_start=1
 /product="death-associated kinase 2"
 /protein_id="AA022165.1"
 /db_xref="GI:18381097"
 /db_xref="locusid:13143"
 /translation="MVOASMRSPNMFPRKQKVEDFYDIGEELGGGPAIVKRCRKS
 TGLVYAKRIKKRQSRASRGVSRREIEREVNITLHDIPEKRTDVLLLEVSGLP
 DFLAKESITDEDAEPLKQILDGVHLHSHKRAHFDKPNIMLDKNVNPRIKLI
 DGIAMKIEAGNEFNIPTGPEVAPETVAYEPLGLADMSIGVITILLGASPL
 GETTQETLNTSAVYDDEEYFNTSELAQDFRLVQPKPKMTIAQSLSEHWK
 AIRRNVDGSGCRKPERRLKTPKTKYTLKSHSLPPNNSYADPFERSSVLEAA
 AEEGRLRECRSRICHDVEYALAAIYEKEMVRESLSQDLRLAQELTKTEAK
 RQAEPAKGLITSGLRKFRSLNRYEALAKQVASEKRFQDLVRALEQKQGYE
 CGLRL"

ORIGIN

Query Match 2.4%; Score 26; DB 10; Length 1883;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 445 CACCCCAACATCATCAGCTGCACGA 470
 DB 337 CACCCCAACATCATCAGCTGCACGA 362

RESULT 18 AB022341 2055 bp mRNA linear PRI 20-APR-2002
 LOCUS Homo sapiens mRNA for ZIP kinase, complete cds.
 DEFINITION AB022341
 ACCESSION AB022341.1 GI:5162863
 VERSION ZIP kinase.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 Murata-Hori, M., Suizu, F., Iwasaki, T., Kikuchi, A. and Hosoya, H.
 ZIP kinase identified as a novel myosin regulatory light chain
 kinase in HeLa cells
 FEBS Lett. 451 (1), 81-84 (1999)
 JOURNAL MEDLINE
 PUBMED 10356987
 2 (bases 1 to 2055)
 AUTHOR Iwasaki, T., Murata-Hori, M. and Hosoya, H.
 TITLE Direct Submission

JOURNAL

Submitted (12-JAN-1999) Hiroshi Hosoya, Faculty of Science, Hiroshima University, Department of Biological Science; Kugamiyama 1-3-1, Higashi-Hiroshima, Hiroshima 739-8526, Japan (E-mail: hohosoya@sci.hiroshima-u.ac.jp, Tel: 81-824-24-7443, Fax: 81-824-24-0734)

FEATURES

Location/Qualifiers
1..2055

source
/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_line="Hela RCB0007"

/clone_lib="Hela cell cDNA library"

1..2055

/gene="ZIPK"

46..1410

/gene="ZIPK"

/product="ZIP kinase"

/protein_id="BAA81746.1"

/db_xref="GI:5162884"

/translation="MSTFRQEDVEDHYENGEEIGSGQPAIVKCKQKGTKEVAAKFI
KKRLSSSRGVREIEREVNLRIRHNPITLHDIFENKIDVLLILEVSGSELF
DPLAKESLTEDATQLOKIDGVYHLSKRIAHFDLPENIMLDKVPNPKIL
DPEIAKIEAGNEFNKI FGPPEFVAPETVVEPLGLADWMSIGVITVLLSGAPPL
GETKQETLVNISAVNDPDESEFNTSELAKDPIRLVADKPRRMTIAQSHSMIX
AIRRRVGRGDSGRKPRRRKTTKRLKVTYSKSHSLPPNMSYADPEPRSKVLEAA
AEEGLREIQRSLRCHDEYBALAAIYEKEMVRESDSLQDURLRDELKTALK
KQAEKRGALGTSGLKRFRLNRYBALAKQVASEKRFVQDURLVRLDEKLGVE
GSLR"

2055

/gene="ZIPK"

/note="24 a nucleotide"

polya_site

ORIGIN

Query Match 2.4%; Score 26; DB 9; Length 2055;
Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CAGTTCCTCAAGCAGATCCTGACGG 604

Db 397 CAGTTCCTCAAGCAGATCCTGACGG 422

RESULT 19

BD127287 2079 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD127287 2079 bp DNA linear PAT 18-SEP-2002
Primer for synthesizing full-length cDNA and use thereof.
BD127287
UP 2002017375-A/2718
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2079)
Oca.T., Nishikawa.T., Isogai.T., Hayashi.K., Ishii.S., Kawai.Y.,
Wakamatsu.A., Sugiyama.T., Nagai.K., Kojima.S., Otsuki.T. and
Koga.H.

Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 2718 22-JAN-2002;
HELIX RESEARCH INSTITUTE

COMMENT
OS Homo sapiens (human)
PN 22-JAN-2002
PD 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOTAMA,
PI TETSUO OTSUKI, HISASHI KOGA

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,

FEATURES

source

Location/Qualifiers
1..2079

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2079;
Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CAGTTCCTCAAGCAGATCCTGACGG 604

Db 421 CAGTTCCTCAAGCAGATCCTGACGG 446

RESULT 20

AK074799 2079 bp mRNA linear PRI 03-SEP-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK074799 2079 bp mRNA linear PRI 03-SEP-2002
Homo sapiens cDNA FLJ9318 f1s, clone NT28P2001529, highly similar
to Homo sapiens mRNA for ZIP-kinase.
AK074799.1 GI:22760480
oligo capping: f1s (full insert sequence).
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1

Isogai.T., Oca.T., Nishikawa.T., Hayashi.K., Otsuki.T.,
Sugiyama.T., Suzuki.Y., Nagai.K., Sugano.S., Ishii.S.,
Kawai.Hio.Y., Saito.K., Yamamoto.J., Wakamatsu.A., Nakamura.Y.,
Kojima.S., Nagahara.K., Masuno.Y., Ono.T., Okano.K., Yoshikawa.Y.,
Aotsuka.S., Sasaki.N., Hattori.A., Okumura.K., Iwayanagi.T. and
Ninomiya.K.

NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 2079)
Isogai.T. and Otsuki.T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1533-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1..2079

FEATURES

source

Location/Qualifiers
1..2079

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT28P2001529"

/cell_line="NT2"

/cell_type="leucocarcinoma"

/clone_lib="NT2RP2"

/note="cloning vector: pMD18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"

ORIGIN

Query Match 2.4%; Score 26; DB 9; Length 2079;
Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CAGTTCCTCAAGCAGATCCTGACGG 604

Db 421 CAGTTCCTCAGCAGATCCTCGACGG 446

RESULT 21
BD176607

LOCUS BD176607 2105 bp DNA linear PAT 18-MAR-2003

DEFINITION Method of molecular diagnosis of chronic myelogenous leukemia.

ACCESSION BD176607

VERSION BD176607.1 GI:29122317

KEYWORDS WO 02070747-A/9.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 1 (bases 1 to 2105)

COMMENT
Method of molecular diagnosis of chronic myelogenous leukemia
Patent: WO 02070747-A 9 12-SEP-2002;
FUJISAWA PHARMACEUTICAL CO LTD, HIROYUKI MAMO
OS Homo sapiens (human)
PN WO 02070747-A/9
PD 12-SEP-2002
PF 01-MAR-2002 WO 2002JP001901
PR 01-MAR-2001 JP 01P 056438
PI HIROYUKI MAMO
PC C1Q01/68, C12Q01/02, G01N33/15, C12N15/09, A61P35/02 CC Method of
molecular diagnosis of chronic myelogenous leukemia FH Key
Location/Qualifiers
FT CDS (94).. (1458).

FEATURES
source 1..2105
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2105;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAGCAGATCCTCGACGG 604

Db 445 CAGTTCCTCAGCAGATCCTCGACGG 470

RESULT 22
AB007144

LOCUS AB007144 2105 bp mRNA linear PRI 25-FEB-1998

DEFINITION Homo sapiens mRNA for ZIP-kinase, complete cds.

ACCESSION AB007144

VERSION AB007144.1 GI:2911155

KEYWORDS ZIP-kinase.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
AUTHORS Kawai,T., Matsumoto,M., Takeda,K., Sanjo,H. and Akira,S.
TITLE 1 (sites)
JOURNAL ZIP kinase, a novel serine/threonine kinase which mediates
MEDLINE apoptosis
PUBMED 948481
948481
2 (bases 1 to 2105)
AUTHORS Akira,S. and Kawai,T.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1997) Shinzo Akira, Hyogo College of Medicine,
Department of Biochemistry, Mukogawa-cho 1-1, Nishinomiya, Hyogo
663, Japan (E-mail:akira@hyo-med.ac.jp, Tel:+81-798-45-6357,
Fax:+81-798-46-3164)
Location/Qualifiers
1..2105

CDS

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
94..1458
/codon_start=1
/product="ZIP-kinase"
/protein_id="BAA24955.1"
/db_xref="GI:2911156"

polyA_site
2105
/note="27 a nucleotides"

ORIGIN

Query Match 2.4%; Score 26; DB 9; Length 2105;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAGCAGATCCTCGACGG 604

Db 445 CAGTTCCTCAGCAGATCCTCGACGG 470

RESULT 23
AK097643

LOCUS AK097643 2118 bp mRNA linear PRI 15-JUL-2002

DEFINITION Homo sapiens cDNA FLJ40324 fis, clone TEST12031006, highly similar
to Homo sapiens mRNA for ZIP-kinase.

ACCESSION AK097643

VERSION AK097643.1 GI:21757480

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 1
Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,
Tanai,H., Watanabe,S., Ishida,S., Oho,Y., Hotuta,T., Watanabe,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K.
and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2118)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kasarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST12031006"

/c1sue_type="testis"
/clone_id="TEST12"
/note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 2.4%; Score 26; DB 9; Length 2118;
Best Local Similarity 100.0%; Pred. No. 0.028; 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGCAGATCTCGACGG 604
|||||
DQ 460 CAGTTCTCAAGCAGATCTCGACGG 485

RESULT 24
AR076189 2132 bp DNA linear PAT 30-AUG-2000
LOCUS AR076189
DEFINITION Sequence 3 from patent US 5958748.
ACCESSION AR076189
VERSION AR076189.1 GI:10002935
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2132)
AUTHORS Akira,S. and Kawai,T.
TITLE DNA coding for serine/threonine kinase
JOURNAL Patent: US 5958748-A 3 28-SEP-1999;
FEATURES
Location/Qualifiers
1..2132
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.028; 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGCAGATCTCGACGG 604
|||||
DQ 445 CAGTTCTCAAGCAGATCTCGACGG 470

RESULT 25
AR124102 2132 bp DNA linear PAT 16-MAY-2001
LOCUS AR124102
DEFINITION Sequence 3 from patent US 6171841.
ACCESSION AR124102
VERSION AR124102.1 GI:14109463
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2132)
AUTHORS Akira,S. and Kawai,T.
TITLE DNA coding for serine/threonine kinase
JOURNAL Patent: US 6171841-A 3 09-JAN-2001;
FEATURES
Location/Qualifiers
1..2132
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.028; 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGCAGATCTCGACGG 604
|||||
DQ 445 CAGTTCTCAAGCAGATCTCGACGG 470

RESULT 26

E23384 2132 bp DNA linear PAT 18-JUN-2001
LOCUS E23384
DEFINITION DNA encoding serine/threonine kinase.
ACCESSION E23384
VERSION E23384.1 GI:13024386
KEYWORDS JP 199098984-A/1.
SOURCE unidentified
ORGANISM unidentified

REFERENCE 1 (bases 1 to 2132)
AUTHORS Shizuo,S. and Taro,K.
TITLE DNA encoding serine/threonine kinase
JOURNAL Patent: JP 199098984-A 1 13-APR-1999;
SCIENCE & TECH AGENCY
COMMENT OS Unidentified
PN JP 199098984-A/1
PD 13-APR-1999
PF 26-SEP-1997 JP 1997261589
PR

SHIZUO SHINRA,TARO KAWAI
PI C12N15/09,C12N1/21,C12N9/12//C12N15/09,C12R1/91,C12N1/21,
PC C12R1/19)
PC C12N9/12,C12R1/19),C12N15/00,C12N15/00,C12R1/91) CC
Strandedness: Double;
CC Topology: linear;
FH Key Location/Qualifiers
FT CDS 94..1455.

FEATURES
source
1..2132
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.028; 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGCAGATCTCGACGG 604
|||||
DQ 445 CAGTTCTCAAGCAGATCTCGACGG 470

RESULT 27
AX880020 2224 bp DNA linear PAT 17-DEC-2003
LOCUS AX880020
DEFINITION Sequence 14925 from Patent EP1074617.
ACCESSION AX880020
VERSION AX880020.1 GI:40034756
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,D.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 14925 07-FEB-2001;
FEATURES
Location/Qualifiers
1..2224
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
215..1579
/note="unamed protein product"
/codon_start=1
/protein_id="CA890766.1"
/db_xref="GI:40034757"
/translation="MSTPREDVDVHYEMGEELGSGOPATVRCKRQGTGKEVAAKPI
KKRLSSRRGVSRERIEVNLIRIRHNITLHDIPENKTDVYLILAVSGELF
DPLAEKSLTDEATQFLKQILDGVHLSKRIAHFDLKENITLIDKNVNPRIKLI

CDS

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2224;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAAGCAGATCTCGACGG 604
 DB 566 CAGTTCCTCAAGCAGATCTCGACGG 591

DEGIARKIEAGNEFKXNIFGTPPEVAPPEIVNVEPLGLHEADMDMSIGVITYILLSGASPFL
 GETKQETLNI SAVNYDPDEEYFNSNGELAKDFIRLLYVDKPKMTTIAQSLSEHWIK
 ATRRVRGDESGRKRERRLKTRTKETIKSHSLPNNNSKAPFEPSKYLEPAA
 AEEGRELORSRICHEDVEALAIYEEKAYRESDSLQDLRLQCELLKTEALK
 RQAEERKAGALLTSGIKRRFRSLERLYALAKOVASEKRFQDVLVRLAEKRGVE
 CGLR"

RESULT 28
 LOCUS BD012208 2224 bp DNA linear PAT 02-AUG-2002
 DEFINITION Novel genes encoding protein kinase or protein phosphatase.
 ACCESSION BD012208
 VERSION BD012208.1 GI:22092397
 KEYWORDS NO 0109345-A/6
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2224)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K., Otsuki,T.,
 Funahashi,S., Seno,C. and Nezu,J.
 Novel genes encoding protein kinase or protein phosphatase
 Patent: WO 0109345-A 6 08-FEB-2001;
 HELIX RESEARCH INSTITUTE, TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA,
 KOJI HAYASHI, KAORU SAITO, JUNICHI YAMAMOTO, SHIZUKO ISHII, OMOYASU
 SUGIYAMA, AI WAKAMATSU, KEIICHI NAGAI, TETSUJI OTSUKI, SHINICHI
 FUNAHASHI, HIAKI SENO, JUNICHI NEZU

COMMENT OS Homo sapiens (human)
 PN WO 0109345-A/6
 PD 08-FEB-2001
 PR 28-JUL-2000 WO 2000JP005060
 PF 29-JUL-1999 JP 99P 248036, 11-JAN-2000 JP 00P 118776 PR
 02-MAY-2000 JP 00P 183767, 18-OCT-1999 US 60/159590 PR
 17-FEB-2000 US 60/183322
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, PI
 KAORU SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI, SHINICHI FUNAHASHI, CHIAKI SENO,
 PI JUNICHI NEZU
 PC C12N15/54, C12N15/55, C12N9/12, C12N9/16, C12N5/10, C12N1/21 PC
 C12N1/19, C07K16/40
 PC G01N33/53, G01N33/566

FEATURES
 source 1..2224
 Location/Qualifiers
 FT CDS
 Key Location/Qualifiers
 CDS (215)..(1576).
 1..2224
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 2.4%; Score 26; DB 6; Length 2224;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAAGCAGATCTCGACGG 604
 DB 566 CAGTTCCTCAAGCAGATCTCGACGG 591

RESULT 29

BD158150
 LOCUS BD158150 2224 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD158150
 VERSION BD158150.1 GI:27863908
 KEYWORDS JP 2002191363-A/12993.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2224)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K., Otsuki,T.,
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 12993 09-JUL-2002;
 HELIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)
 PN JP 2002191363-A/12993
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/66//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT CDS
 Location/Qualifiers
 (215)..(1576).
 1..2224
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source 1..2224
 Location/Qualifiers
 FT CDS
 Key Location/Qualifiers
 CDS (215)..(1576).
 1..2224
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 2.4%; Score 26; DB 6; Length 2224;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAAGCAGATCTCGACGG 604
 DB 566 CAGTTCCTCAAGCAGATCTCGACGG 591

RESULT 30
 LOCUS AK027590 2224 bp mRNA linear PRI 01-AUG-2002
 DEFINITION Homo sapiens cDNA FJ14684 file, clone NT2RP2004933, highly similar
 to Homo sapiens mRNA for ZIP-kinase.
 ACCESSION AK027590
 VERSION AK027590.1 GI:14042369
 KEYWORDS OLIGO capping; its (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
 Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
 Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
 Wakamatsu,A., Nakamura,Y., Nagahara,K., Maehuo,Y. and Oshima,A.
 NEDO human cDNA sequencing project

REFERENCE 2 (bases 1 to 2224)
 Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomcshri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

TITLE Unpublished
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3' end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES

source

1..2224
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2004933"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RP2"
/note="cloning vector: pMT18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."

ORIGIN

Query Match 2.4%; Score 26; DB 9; Length 2224;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAAGCAGATCCTGACGG 604
DB 566 CAGTTCCTCAAGCAGATCCTGACGG 591

RESULT 31

AC011488/ c 134308 bp DNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens chromosome 19 clone CTB-171N13, complete sequence.

AC011488
AC011488
AC011488.7 GI:22038500
HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 134308)

REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.
1 (bases 1 to 134308)

REFERENCE Direct Submission
Unpublished

REFERENCE Direct Submission
2 (bases 1 to 134308)

REFERENCE Direct Submission
Direct Submission

REFERENCE Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 134308)

REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission

REFERENCE Submitted (01-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Aug 1, 2002 this sequence version replaced gi:9256301.

REFERENCE Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov

REFERENCE Finishing Completed at Stanford Human Genome Center
www.hgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

FEATURES

source

1..134308
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-171N13"

ORIGIN

Query Match 2.4%; Score 26; DB 9; Length 134308;
Best Local Similarity 100.0%; Pred. No. 0.02;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAAGCAGATCCTGACGG 604
DB 114273 CAGTTCCTCAAGCAGATCCTGACGG 114248

RESULT 32

AC124439 162530 bp DNA linear ROD 13-NOV-2003
LOCUS AC124439
DEFINITION Mus musculus BAC clone RP24-261N14 from chromosome 9, complete
sequence.

AC124439
AC124439.4 GI:28626894
HTG.

REFERENCE Mus musculus (house mouse)
Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 162530)
Bershoff, A., Cotton, M., Meyer, R. and Bielicki, L.
The sequence of Mus musculus BAC clone RP24-261N14

REFERENCE 2 (bases 1 to 162530)
Unpublished (2001)

REFERENCE Wilson, R.
Sequencing of Mus musculus

REFERENCE Direct Submission
Unpublished (2001)

REFERENCE 3 (bases 1 to 162530)
McPherson, J.D. and Waterston, R.H.

REFERENCE Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 162530)

REFERENCE McPherson, J.D. and Waterston, R.H.
Direct Submission

REFERENCE Submitted (16-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 162530)

REFERENCE McPherson, J.D. and Waterston, R.H.
Direct Submission

REFERENCE Submitted (01-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 162530)

REFERENCE Wilson, R.
Sequencing of Mus musculus

REFERENCE Submitted (13-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 1, 2003 this sequence version replaced gi:2764792.

REFERENCE Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC

REFERENCE Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu

REFERENCE Summary Statistics
Center Project name: W_B80261N14

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For

additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCT-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

```
source
1..162530
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/map="9"
/clone="RP24-261N14"
/clone_1fb="RPCT-24"
repeat_region
822..969
/rpt_family="L2"
repeat_region
1071..1215
/rpt_family="Alu"
repeat_region
1358..1482
/rpt_family="Alu"
repeat_region
3386..3436
/rpt_family="Alu"
repeat_region
4087..4252
/rpt_family="ID"
repeat_region
4301..4469
/rpt_family="B4"
repeat_region
4825..4872
/rpt_family="B2"
repeat_region
5607..5671
/rpt_family="Alu"
repeat_region
5736..5850
/rpt_family="Alu"
repeat_region
5928..6139
/rpt_family="Alu"
repeat_region
6140..6226
/rpt_family="B2"
repeat_region
6289..6434
/rpt_family="MER1_type"
repeat_region
6553..6690
/rpt_family="B4"
repeat_region
6745..6886
/rpt_family="Alu"
repeat_region
7013..7056
/rpt_family="Alu"
repeat_region
7689..7833
/rpt_family="Alu"
repeat_region
7872..8045
/rpt_family="Alu"
repeat_region
8406..8469
/rpt_family="B2"
repeat_region
8588..8705
/rpt_family="B2"
repeat_region
9193..9303
/rpt_family="Alu"
repeat_region
9306..9326
/rpt_family="L1"
repeat_region
9327..9418
/rpt_family="B4"
repeat_region
9567..9782
/rpt_family="Alu"
repeat_region
9798..9981
/rpt_family="L1"
repeat_region
10168..10291
/rpt_family="B2"
repeat_region
10293..10485
/rpt_family="Alu"
repeat_region
10562..10711
/rpt_family="L1"
```

```
repeat_region
10832..10970
/rpt_family="Alu"
repeat_region
10986..11005
/rpt_family="B4"
repeat_region
11006..11126
/rpt_family="ERVK"
repeat_region
11127..11252
/rpt_family="Alu"
repeat_region
12519..12657
/rpt_family="ERVK"
repeat_region
12953..13150
/rpt_family="Alu"
repeat_region
13183..13252
/rpt_family="ID"
repeat_region
13254..13392
/rpt_family="B2"
repeat_region
13615..13751
/rpt_family="Alu"
repeat_region
14166..14313
/rpt_family="B2"
repeat_region
14773..14983
/rpt_family="Alu"
repeat_region
15017..15082
/rpt_family="B2"
repeat_region
15213..15553
/rpt_family="ID"
repeat_region
15668..15764
/rpt_family="MALR"
repeat_region
15883..16118
/rpt_family="MER1_type"
repeat_region
16341..16539
/rpt_family="B2"
repeat_region
16404..16584
/rpt_family="MER1_type"
repeat_region
16592..16661
/rpt_family="Y-chromosomal"
repeat_region
16663..16685
/rpt_family="ERV1"
repeat_region
16686..16873
/rpt_family="B4"
repeat_region
16874..16884
/rpt_family="B2"
repeat_region
16885..17032
/rpt_family="B4"
repeat_region
17033..17048
/rpt_family="Alu"
repeat_region
17060..17144
/rpt_family="B4"
repeat_region
17460..17627
/rpt_family="Alu"
repeat_region
17667..17745
/rpt_family="Alu"
repeat_region
17793..18141
/rpt_family="ID"
repeat_region
18233..18260
/rpt_family="MALR"
repeat_region
18261..18333
/rpt_family="Alu"
repeat_region
18334..18419
/rpt_family="ID"
repeat_region
18468..18581
/rpt_family="Alu"
repeat_region
/rpt_family="Alu"
```

Query Match 2.4%; Score 26; DB 10; Length 162530;
Best Local Similarity 100.0%; Pred.No. 0.02;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 445 CACCCACATCATCAGCTGCAGCA 470
DB 96302 CACCCACATCATCAGCTGCAGCA 96327

```
RESULT 33
AC015914      170078 bp      DNA      linear      PRI 28-JUN-2002
DEFINITION    Homo sapiens chromosome 15, clone RP11-111E14, complete sequence.
ACCESSION     AC015914
VERSION       AC015914.8      GI:21622757
KEYWORDS      HTG.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 170078)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-111E14
Unpublished
2 (bases 1 to 170078)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Bouhgailler,B.,
Brown,A., Castelle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Darelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,D.,
Lehocky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Mortom,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 170078)
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhgailler,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Choepe,I., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Darelano,K., Dewar,K., Diaz,U.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Labrecque,K., Lamazares,R.,
Lander,E., Lehocky,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,V., Marquis,N., Matthews,C., McCaithy,M.,
McEwan,P., McKernan,K., Meldrum,J., Menus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Toplam,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.,J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 28, 2002 this sequence version replaced gi:18482289.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/xv/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L839
Center clone name: 111_E_14
----- Location/Qualifiers
source
1. 170078
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-111E14"
/clone_lib="RPCT-11 Human Male BAC"
1. 5923
/rpt_family="L1PA5"
repeat_region
5914..6393
/rpt_family="L1MA3"
repeat_region
6394..6689
/rpt_family="AluY"
repeat_region
6690..6926
/rpt_family="L1MA3"
repeat_region
6933..7412
/rpt_family="L1MA3"
repeat_region
7413..7699
/rpt_family="AluYb"
repeat_region
7700..9996
/rpt_family="L1MA3"
repeat_region
9988..10287
/rpt_family="L1P91"
repeat_region
10293..10619
/rpt_family="L1MA3"
repeat_region
10619..11025
/rpt_family="L1MA3"
repeat_region
11025..11322
/rpt_family="AluX"
repeat_region
11403..11445
/rpt_family="TGAA)n"
repeat_region
11535..11644
/rpt_family="L2"
repeat_region
12418..12568
/rpt_family="MIR"
repeat_region
12676..12703
/rpt_family="AT_rich"
repeat_region
12763..13055
/rpt_family="AluTo"
repeat_region
13090..13118
/rpt_family="AT_rich"
repeat_region
13428..13540
/rpt_family="MIR"
repeat_region
14286..14479
/rpt_family="MIR"
repeat_region
14642..14980
/rpt_family="MIR"
repeat_region
14981..15290
/rpt_family="AluX"
repeat_region
15291..15332
/rpt_family="MIR"
repeat_region
15332..15592
/rpt_family="MIR"
repeat_region
15592..15960
/rpt_family="L2"
repeat_region
16571..16665
/rpt_family="MIR3"
repeat_region
16754..16909
/rpt_family="MIR"
repeat_region
17050..17667
/rpt_family="MIR3"
repeat_region
18126..18215
/rpt_family="FLM_A"
repeat_region
18595..18770
/rpt_family="MIR"
repeat_region
19383..19689
/rpt_family="AluX"
repeat_region
19740..19833
/rpt_family="MIR"
repeat_region
19834..20138
/rpt_family="AluYo"
repeat_region
20139..20281
/rpt_family="MIR"
repeat_region
20303..20593
/rpt_family="MIR"
repeat_region
```

FEATURES

Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Vei,R., Vo,A., Wilson,B., Wu,X., Wymann,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (12-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Dec 12, 2001 this sequence version replaced gi:15983568.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L5589

Center clone name: 505_L_16

FEATURES

Source

Location/Qualifiers

1. 194023

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-505L16"

/clone_1b="RP11 Human Male BAC"

1137. 1432

/rpt_family="AluSx"

complement(2167. 2474)

/rpt_family="AluUo"

2555. 2866

/rpt_family="AluSx"

complement(2901. 3050)

/rpt_family="MIR"

4311. 4374

/rpt_family="L2"

4357. 4491

/rpt_family="MIR"

4385. 4495

/rpt_family="L1MB6"

5969. 6027

/rpt_family="GA-rich"

6182. 6495

/rpt_family="L2"

6573. 6775

/rpt_family="MER2"

8497. 8697

/rpt_family="L1MC5"

8698. 9010

/rpt_family="AluSg"

9011. 9212

/rpt_family="L1MC5"

10179. 10458

/rpt_family="AluSx"

complement(11093. 11402)

/rpt_family="AluSx"

11655. 11749

complement(11770. 12035)

/rpt_family="L1MED4"

12097. 12164

/rpt_family="MER104"

12559. 12836

/rpt_family="L2"

12899. 13048

/rpt_family="MIR"

13677. 13854

/rpt_family="MIR"

complement(15473. 15602)

/rpt_family="FLAM_C"

17007. 17057

/rpt_family="MER115"

17296. 17339

/rpt_family="(CA)n"

18443. 18585

/rpt_family="L2"

18583. 18623

/rpt_family="MIR"

19151. 19319

/rpt_family="L2"

complement(19361. 19641)

/rpt_family="AluUo"

complement(19740. 19894)

/rpt_family="MER5B"

complement(20150. 20326)

/rpt_family="MIR"

20403. 20479

/rpt_family="GA-rich"

20538. 20831

/rpt_family="AluSg"

21384. 21568

/rpt_family="MIR"

21623. 21698

/rpt_family="L2"

22239. 22393

/rpt_family="MLT1A1"

22354. 22706

/rpt_family="AluUo"

22707. 22820

/rpt_family="MLT1A1"

23504. 23800

/rpt_family="AluSx"

23801. 24111

/rpt_family="AluSg"

24148. 24177

/rpt_family="AT-rich"

complement(24399. 24678)

/rpt_family="L1MB3"

complement(24851. 25159)

/rpt_family="AluSg"

complement(25214. 25511)

/rpt_family="AluUo"

25615. 25664

/rpt_family="T-rich"

complement(25867. 26173)

/rpt_family="AluSx"

complement(26176. 26476)

/rpt_family="AluSx"

Query Match 2.4%; Score 26; DB 9; Length 194023;

Best local Similarity 100.0%; Pred. No. 0.02;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 923 GGCGTCATCACCTACCTCGTGTAGT 948

DB 21187 GGCGTCATCACCTACCTCGTGTAGT 21162

RESULT 35

AC034201/c

LOCUS

DEFINITION Homo sapiens chromosome 19 clone CIB-172H21, WORKING DRAFT

SEQUENCE 35 unordered pieces.

ACCESSION AC034201

VERSION AC034201.2 GI:12965321

KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 210617)

AUTHORS DOE Joint Genome Institute.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sequencing of Human Chromosome 19
2 (bases 1 to 210617)
DOE Joint Genome Institute.
Direct Submission
Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Feb 20, 2001 this sequence version replaced gi:7417678.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 130441, BC95464
Center clone name: CIT978SKB_172H21

Summary Statistics
Consensus quality: 178878 bases at least Q40
Consensus quality: 191077 bases at least Q30
Consensus quality: 196408 bases at least Q20
Estimated insert size: 138740; agarose-fp estimation
Estimated insert size: 207217; sum-of-contigs estimation
Quality coverage: 5.87 in Q20 bases; agarose-fp estimation
Quality coverage: 8.77 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1085: contig of 1085 bp in length
* 1086 1185: gap of unknown length
* 1186 2364: contig of 1179 bp in length
* 2365 2464: gap of unknown length
* 2465 3612: contig of 1148 bp in length
* 3613 3712: gap of unknown length
* 3713 4856: contig of 1144 bp in length
* 4857 4956: gap of unknown length
* 4957 6069: contig of 1113 bp in length
* 6070 6169: gap of unknown length
* 6170 7483: contig of 1314 bp in length
* 7484 7583: gap of unknown length
* 7584 9141: contig of 1558 bp in length
* 9142 9241: gap of unknown length
* 9242 11367: contig of 2126 bp in length
* 11368 11467: gap of unknown length
* 11469 12787: contig of 1320 bp in length
* 12788 12887: gap of unknown length
* 12889 14416: contig of 1523 bp in length
* 14417 14516: gap of unknown length
* 14517 16383: contig of 1867 bp in length
* 16384 16483: gap of unknown length
* 16484 18681: contig of 2198 bp in length
* 18682 18781: gap of unknown length
* 18782 20834: contig of 2053 bp in length
* 20835 20934: gap of unknown length
* 20935 22101: contig of 1167 bp in length
* 22102 22201: gap of unknown length
* 22202 24304: contig of 2103 bp in length
* 24305 24404: gap of unknown length
* 24406 26002: contig of 1598 bp in length
* 26003 26102: gap of unknown length
* 26103 28540: contig of 2438 bp in length
* 28541 28640: gap of unknown length
* 28641 32193: contig of 3353 bp in length
* 32194 32293: gap of unknown length
* 32294 34163: contig of 1870 bp in length
* 34164 34263: gap of unknown length
* 34264 37660: contig of 3397 bp in length
* 37661 37760: gap of unknown length

37761 40553: contig of 2793 bp in length
* 40554 40653: gap of unknown length
* 40654 44170: contig of 3517 bp in length
* 44171 44270: gap of unknown length
* 44271 47648: contig of 3378 bp in length
* 47649 47748: gap of unknown length
* 47749 52104: contig of 4356 bp in length
* 52105 52204: gap of unknown length
* 52205 54768: contig of 2564 bp in length
* 54769 54868: gap of unknown length
* 54869 62590: contig of 7722 bp in length
* 62591 62690: gap of unknown length
* 62691 71952: contig of 9262 bp in length
* 71953 72052: gap of unknown length
* 72053 82465: contig of 10413 bp in length
* 82466 82565: gap of unknown length
* 82566 87764: contig of 5199 bp in length
* 87765 87864: gap of unknown length
* 87865 95981: contig of 8117 bp in length
* 95982 96081: gap of unknown length
* 96082 108564: contig of 12483 bp in length
* 108565 108664: gap of unknown length
* 108666 123979: contig of 15315 bp in length
* 123980 124079: gap of unknown length
* 124080 140602: contig of 16523 bp in length
* 140603 140702: gap of unknown length
* 140703 174140: contig of 33438 bp in length
* 174141 174240: gap of unknown length
* 174241 210617: contig of 36377 bp in length.

FEATURES

source

1..210617
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-172H21"
/clone_lib="Caltech human BAC library B"

ORIGIN

Query Match 2.4%; Score 26; DB 2; Length 210617;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 579 CAGTTCCTCAGCAGATCCTGACG 604
DB 117318 CAGTTCCTCAGCAGATCCTGACG 117293

RESULT 36

AC016179

LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-652K13 map 15, WORKING DRAFT
AC016179
AC016179.6 GI:14336574
ACCESSION
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 224148)
Britten, B., Linton, L., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-652K13
unpublished
2 (bases 1 to 224148)
Britten, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, V., Barna, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collins, A.,
Cooke, P., DeArrellano, K., Dewar, K., Domingo, M., Dorelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardina, S., Grant, G., Hagos, B., Harford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lenockky, J., Lieu, C., Locke, K., Macdonald, P., Marcus, N.,

TITLE
JOURNAL

COMMENT

McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Petersen, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testa, S., Titrill, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission

Submitted (23-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2001 this sequence version replaced gi:13621297.
All repeats were identified using RepeatMasker:
Shit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4221

Center clone name: 652_K.13

----- Summary Statistics

Sequencing vector: M13; M7815; 3% of reads

Chemistry: Dye-terminator Big Dye, 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 204388 bases at least Q40

Consensus quality: 212270 bases at least Q30

Consensus quality: 215891 bases at least Q20

Insert size: 194000; agarose-fp

Insert size: 220248; sum-of-contigs

Quality coverage: 6.5 in Q20 bases; agarose-fp

Quality coverage: 5.7 in Q20 ba.

NOTE: This is a 'working draft' sequence. It currently
consists of 40 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```

* 1 991: contig of 991 bp in length
* 992 1091: gap of 100 bp
* 1092 1700: contig of 609 bp in length
* 1701 1800: gap of 100 bp
* 1801 2218: contig of 418 bp in length
* 2219 2318: gap of 100 bp
* 2319 2758: contig of 440 bp in length
* 2759 2858: gap of 100 bp
* 2859 3289: contig of 431 bp in length
* 3290 3389: gap of 100 bp
* 3390 4308: contig of 920 bp in length
* 4310 4409: gap of 100 bp
* 4410 4916: contig of 507 bp in length
* 4917 5016: gap of 100 bp
* 5017 5511: contig of 495 bp in length
* 5512 5611: gap of 100 bp
* 5612 6450: contig of 839 bp in length
* 6451 6550: gap of 100 bp
* 6551 7042: contig of 492 bp in length
* 7043 7142: gap of 100 bp
* 7143 7959: contig of 817 bp in length
* 7960 8059: gap of 100 bp
* 8060 8731: contig of 672 bp in length
* 8732 8831: gap of 100 bp
* 8833 9608: contig of 777 bp in length
* 9609 9708: gap of 100 bp
* 9709 10418: contig of 710 bp in length
* 10419 10518: gap of 100 bp
* 10519 11206: contig of 688 bp in length
* 11207 11306: gap of 100 bp
* 11307 11908: contig of 602 bp in length
* 11909 12008: gap of 100 bp

```

FEATURES

source

```

* 12009 12796: contig of 788 bp in length
* 12797 12896: gap of 100 bp
* 12897 13368: contig of 472 bp in length
* 13369 13468: gap of 100 bp
* 13469 13308: contig of 841 bp in length
* 13310 14409: gap of 100 bp
* 14410 15805: contig of 1396 bp in length
* 15806 15905: gap of 100 bp
* 15906 16715: contig of 810 bp in length
* 16716 16815: gap of 100 bp
* 16816 17848: contig of 1033 bp in length
* 17849 17948: gap of 100 bp
* 17949 19466: contig of 1518 bp in length
* 19467 19566: gap of 100 bp
* 19567 20377: contig of 811 bp in length
* 20378 20478: gap of 100 bp
* 20479 21591: contig of 1113 bp in length
* 21592 21690: gap of 100 bp
* 21691 22485: contig of 795 bp in length
* 22486 22585: gap of 100 bp
* 22586 23555: contig of 970 bp in length
* 23556 23655: gap of 100 bp
* 23656 26404: contig of 2749 bp in length
* 26405 26504: gap of 100 bp
* 26505 26504: contig of 5303 bp in length
* 26506 27936: contig of 2733 bp in length
* 27937 28388: gap of 100 bp
* 28389 82489: gap of 100 bp
* 82490 87677: contig of 5278 bp in length
* 87678 87768: gap of 100 bp
* 87769 93341: contig of 5474 bp in length
* 93342 93441: gap of 100 bp
* 93442 103488: contig of 10048 bp in length
* 103489 103589: gap of 100 bp
* 103590 117998: contig of 14409 bp in length
* 117999 118098: gap of 100 bp
* 118099 133427: contig of 15329 bp in length
* 133428 133527: gap of 100 bp
* 133528 148245: contig of 15718 bp in length
* 148246 149246: gap of 100 bp
* 149247 162652: contig of 13307 bp in length
* 162653 162752: gap of 100 bp
* 162753 188495: contig of 25743 bp in length
* 188496 188595: gap of 100 bp
* 188596 223243: contig of 34648 bp in length
* 223244 223344: gap of 100 bp
* 223345 224148: contig of 805 bp in length.
* 224149 224148: Location/Qualifiers
* 1. 224148
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="15"
  /map="15"
  /clone="RP11-652K13"
  /clone_id="RPC1-11 Human Male BAC"
  1. 991
    /note="assembly_fragment"
    1092..1700
      /note="assembly_fragment"
      1801..2218
        /note="assembly_fragment"
        2319..2758
          /note="assembly_fragment"
          2859..3289
            /note="assembly_fragment"
            3390..4309
              /note="assembly_fragment"
              4410..4916
                /note="assembly_fragment"
                5017..5511
                  /note="assembly_fragment"
                  5612..6450

```

```

misc_feature /note="assembly_fragment"
6551..7042
/note="assembly_fragment"
7143..7359
/note="assembly_fragment"
8060..8731
/note="assembly_fragment"
8832..9608
/note="assembly_fragment"
9709..10418
/note="assembly_fragment"
10519..11206
/note="assembly_fragment"
11307..11908
/note="assembly_fragment"
12009..12756
/note="assembly_fragment"
12897..13368
/note="assembly_fragment"
13469..14309
/note="assembly_fragment"
14410..15805
/note="assembly_fragment"
15906..16715
/note="assembly_fragment"
16816..17848
/note="assembly_fragment"
17949..19466
/note="assembly_fragment"
19567..20377
/note="assembly_fragment"

```

```

Query Match 2.4% Score 26; DB 2; Length 224148;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 923 GGGCTCATCCTCATCCTCTGTAGT 948
Db 27599 GGGCTCATCCTCATCCTCTGTAGT 27624

```

```

RESULT 37
AC096406 228405 bp DNA linear HTG 10-MAY-2003
LOCUS AC096406
DEFINITION Rattus norvegicus clone CH230-62C22, WORKING DRAFT SEQUENCE.
ACCESSION AC096406
VERSION AC096406.6 GI:30520479
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Eumalaya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 228405)

```

```

REFERENCE
AUTHORS Muzny, D., Matre, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Alyea, E., Altschul, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazo, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C. M., Gadi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hughes, M.,

```

```

TITLE JOURNAL
REFERENCE JOURNAL
AUTHORS JOURNAL
TITLE JOURNAL

```

```

REFERENCE JOURNAL
AUTHORS JOURNAL
TITLE JOURNAL

```

COMMENT

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819671.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFHT
Center clone name: CH230-62C22
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 218033 bases at least Q40
Consensus quality: 220284 bases at least Q30
Consensus quality: 221549 bases at least Q20
Estimated insert size: 227542; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

```

Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowls, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu W., Liu Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisedge, H., Lorado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangun, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelern, O., Okunnu, G., Olariunsacon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Plambeck, C., Plopper, F., Poldinger, A., Popovic, D., Primm, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, X., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 228405 contig of 228405 bp in length.
 Location/Qualifiers

FEATURES

source
 1..228405
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-62C22"
 1..1295
 /note="wgs end extension
 clone_end:Sp6"
 4688..5551
 /note="clone boundary
 clone_end:Sp6"
 site:ECORI
 end_sequence:BH287208"
 complement(223054..223821)
 /note="clone boundary
 clone_end:T7
 site:ECORI
 end_sequence:BH287201"
 225678..228405
 /note="wgs end extension
 clone_end:T7"

ORIGIN

Query Match 2.4%; Score 26; DB 2; Length 228405;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 445 CACCCCAACATCATCATGCTGACGCA 470
 Db 212866 CACCCCAACATCATCATGCTGACGCA 212861

RESULT 38
 AC103501
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-98D19, *** SEQUENCING IN PROGRESS
 *** 2 unordered pieces.
 AC103501
 AC103501.6 GI:30520572
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 232704)
 Muzny,D,Marler,M, Lee,A, Abramson,S, Adams,C, Alder,J,
 Allen,C, Allen,H, Alshrocks,S, Amin,A, Anguiano,D,
 Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,B, Baden,H,
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
 Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
 Bryant,N, Buhay,C, Burch,P, Burrell,K, Cessari,H, Center,A,
 Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Calderon,B,
 Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
 Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
 Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
 Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
 Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
 Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
 Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Guerra,M.,
 Gebregeorgis,E., Geier,K., Gill,R., Grady,M., Guerra,T., Guevara,W.,
 Gunaratne,P., Haaland,M., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Huily,K.S., Hume,J., Iglehart,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Khan,Z., Khatib,R., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,B.,
 Lorenshaw,L., Louisgied,H., Lozard,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenan,E.,
 Mijovic,I., Minner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
 Nwakoelamleh,O., Okwou,G., Olarinmoye,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Polidexter,A., Popovic,D., Primus,E., Pu,L.,
 Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reish,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Saverly,G., Scherer,S., Scott,C., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosé,J.,
 Steimle,M., Strong,R., Sutton,A., Swalek,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos-Z, Usmani,K.,
 Valas,R., Vera,V., Villasa,D., Waldron,L., Walker,B., Wang,D.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Weinstock,G. and Gibbs,R.A., Smith,D.R., Holt,R.A., Smith,H.O.,
 Direct Submission
 Unpublished
 2 (bases 1 to 232704)
 Worley,K.C.
 Direct Submission
 Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 232704)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GHTS
 Center clone name: CH230-98D19
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 215152 bases at least Q40

Consensus quality: 217830 bases at least Q30
 Consensus quality: 219818 bases at least Q20
 Estimated insert size: 228227; sum-of-coverage estimation
 Quality coverage: 6x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 2 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 119953: contig of 119953 bp in length
 * 119954 120053: gap of unknown length
 * 120054 232704: contig of 112651 bp in length.
 Location/Qualifiers

source
 1. 232704
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-98D19"
 1. 1508
 /note="wgs_contig"
 5596. 7623
 /note="wgs_contig"

misc_feature
 misc_feature
 misc_feature
 /note="wgs_contig"
 5596. 7623
 /note="wgs_contig"

ORIGIN

Query Match 2.4%; Score 26; DB 2; Length 232704;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 445 CACCCACATCATCGCTGACGCA 470
 Db 151367 CACCCACATCATCGCTGACGCA 151392

RESULT 39
 AC129021 184961 bp DNA linear ROD 08-NOV-2003
 LOCUS Mus musculus BAC clone RP23-312H15 from 16, complete sequence.
 DEFINITION AC129021
 AC129021 GI:27657604
 VERSION HTG.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 184961)
 Belter, E., Haakenson, W. and Kozlowski, A.
 The sequence of Mus musculus BAC clone RP23-312H15
 Unpublished (2001)
 2 (bases 1 to 184961)
 Wilson, R.
 Sequencing of Mus musculus
 Unpublished (2001)
 3 (bases 1 to 184961)
 McPherson, J.D. and Waterston, R.H.
 Direct Submision
 Submitted (24-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 184961)
 McPherson, J.D. and Waterston, R.H.
 Direct Submision
 Submitted (21-DEC-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 184961)
 McPherson, J.D. and Waterston, R.H.
 Direct Submision
 Submitted (12-JAN-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

6 (bases 1 to 184961)
 McPherson, J.D. and Waterston, R.H.
 Direct Submision
 Submitted (25-MAR-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 7 (bases 1 to 184961)
 Wilson, R.
 Direct Submision
 Submitted (08-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jan 12, 2003 this sequence version replaced gi:27356760.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: M_BA0312H15

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
 The RBC1-23 BAC Library has been constructed by Kazutoyo Oseigawa
 and Minako Tatemio in the laboratory of Pieter de Jong
 (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
 brain genomic DNA. The clone and detailed information can be
 obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
 Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
 This sequence is the entire insert of the clone. This clone is
 overlapped by AC087900.
 Location/Qualifiers

1. 184961
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="16"
 /map="16"
 /clone="RP23-312H15"
 /clone_1b="RBC1-23"
 1687. 1826
 /rpt_family="Alu"
 1831. 2015
 /rpt_family="B2"
 2029. 2131
 /rpt_family="Alu"
 2236. 2377
 /rpt_family="MALR"
 2283. 2394
 /rpt_family="B4"
 4004. 4117
 /rpt_family="Alu"
 4150. 4451
 /rpt_family="MALR"

```

repeat_region 4633..4739
/rpt_family="Alu"
repeat_region 4866..4960
/rpt_family="B4"
repeat_region 5006..5152
/rpt_family="Alu"
repeat_region 5216..5314
/rpt_family="B4"
repeat_region 5315..5458
/rpt_family="Alu"
repeat_region 5468..5646
/rpt_family="B4"
repeat_region 5676..5733
/rpt_family="Alu"
repeat_region 5763..6014
/rpt_family="B4"
repeat_region 6252..6450
/rpt_family="B2"
repeat_region 7429..7554
/rpt_family="Alu"
repeat_region 7751..7898
/rpt_family="Alu"
repeat_region 7902..8255
/rpt_family="L2"
repeat_region 8322..8687
/rpt_family="L2"
repeat_region 10163..10220
/rpt_family="MaLR"
repeat_region 10221..10316
/rpt_family="Alu"
repeat_region 10379..10480
/rpt_family="Alu"
repeat_region 10513..10640
/rpt_family="B4"
repeat_region 10679..10754
/rpt_family="ID"
repeat_region 11090..12879
/rpt_family="RMR1A"
repeat_region 13497..13607
/rpt_family="B4"
repeat_region 14650..14801
/rpt_family="B2"
repeat_region 14711..14802
/rpt_family="B4"
repeat_region 14803..14923
/rpt_family="B2"
repeat_region 16125..16271
/rpt_family="Alu"
repeat_region 16354..16723
/rpt_family="L1"
repeat_region 16867..17334
/rpt_family="L1"
repeat_region 17382..17545
/rpt_family="B4"
repeat_region 17917..18130
/rpt_family="B2"
repeat_region 18451..18597
/rpt_family="Alu"
repeat_region 18726..18783
/rpt_family="Alu"
repeat_region 21376..21471
/rpt_family="Alu"
repeat_region 21520..21654
/rpt_family="B2"
repeat_region 21674..21813
/rpt_family="Alu"
repeat_region 21818..22022
/rpt_family="RMR1A"
repeat_region 22169..22263
/rpt_family="U6"
repeat_region 22301..22628
/rpt_family="MaLR"
repeat_region 22711..22858

```

```

repeat_region /rpt_family="Alu"
24854..24995
/rpt_family="Alu"
repeat_region 25262..25330
/rpt_family="ID"
repeat_region 25377..25523
/rpt_family="Alu"
repeat_region 25744..25931
/rpt_family="B2"
repeat_region 26028..26163
/rpt_family="Alu"
repeat_region 26746..26906
/rpt_family="L1"
repeat_region 27102..27407
/rpt_family="L1"
repeat_region 27479..27537
/rpt_family="ID"
repeat_region 27701..27842
/rpt_family="Alu"
repeat_region 28199..28311
/rpt_family="Alu"
repeat_region 28796..28868

```

Query Match 2.3%; Score 25; DB 10; Length 184961;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 495 GGTGCTGATCTTGAGCTGCTGCTC 519
 Db 21148 GTGCTGATCTTGAGCTGCTGCTC 21172

```

RESULT 40
LOCUS AC087900
DEFINITION Mus musculus clone rp23-17116 map 19 strain C57BL/6J, complete
sequence.
ACCESSION AC087900
VERSION AC087900.19 GI:24961320
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195066)
Cox,T., Shauli,S., Yao,Z. and Roe,B.A.
Mus musculus Chromosome 16 BAC Clone rp23-17116
Unpublished
2 (bases 1 to 195066)
Cox,T., Shauli,S., Yao,Z. and Roe,B.A.
Direct Submission
Submitted (06-FEB-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 195066)
Cox,T., Shauli,S., Yao,Z. and Roe,B.A.
Direct Submission
Submitted (01-NOV-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 195066)
Cox,T., Shauli,S., Yao,Z. and Roe,B.A.
Direct Submission
Submitted (01-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 195066)
Cox,T., Shauli,S., Yao,Z. and Roe,B.A.
Direct Submission
Submitted (14-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 195066)
REFERENCE
AUTHORS
TITLE
JOURNAL

```

AUTHORS Cox, T., Shaull, S., Yao, Z. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-2003) Department of

JOURNAL Submitted (18-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Nov 14, 2002 this sequence version replaced gl:22038563.

----- genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES	Location/Qualifiers
source	1. .195066

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/map="19"
/clone="RP23-17116"
/clone_1db="RPCL - 23 Female (C57BL/6J) Mouse BAC Library"

```

Query Match	2.3%;	Score 25;	DB 10;	Length 195066;
Best Local Similarity	100.0%;	Pred. No. 0.07;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

y 495 GGTGCTGATCTTGGAGCTGTGTC 519
 123208 GGTGCTGATCTTGGAGCTGTGTC 123232

	1800 bp	mRNA	linear	ROD 27-APR-1993
RESULT 41				
AT56KIN3				
LOCUS				
REINITION				
COESSION				
VERSION				
KEYWORDS				
CURCE				
ORGANISM				
	RAT56KIN3			
	Rat S6 kinase mRNA, complete cds.			
	M57428 M35864			
	M57428.1 GI:206839			
	S6 kinase.			
	Rattus norvegicus (Norway rat)			
	Rattus norvegicus			
	Rattus norvegicus			

REFERENCE

1 (bases 1 to 1800)

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

Kozma, S.C., Ferrari, S., Bassand, P., Siegmann, M., Totty, N. and

TITLE	JOURNAL
Cloning of the mitogen-activated S6 kinase from rat liver reveals an enzyme of the second messenger subfamily	Proc. Natl. Acad. Sci. U.S.A. 87 (19), 7365-7369 (1990)

MEDLINE	91017506
PUBMED	1699226
REFERENCE	2 (sites)

TITLE Identification and early activation of a *Xenopus laevis* p70s6k following progesterone-induced meiotic maturation
JOURNAL EMBO J. 11 (5), 1743-1749 (1992)
PAGES 6384

COMMENT	Original source text: Rat (strain Sprague-Dawley) liver, cDNA to mRNA.
PUBMED	1374712
MEDLINE	92258384

FEATURES	Location/Qualifiers
source	1. .1800

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
strain="Sorema-Daw[au]"

```

```

/ad_xref="taxon:10116"
/tissue_type="liver"
/dev_stage="7-8 week"
/tissue_1ib="stratagene lambda zan TI 936507"

```

CDS	/gene="S6 kinase"
	134.1642
	/gene="S6 kinase"

```

/codon_start=1
/product="H6_kinase"
/protein_id="AA442103.1"
/db_xref="GI:206840"
/translation="MAGVIEDIDQEDAGSEDELEEGQINESMDHGCVGYPYELIGNE
HCKFELSETSVNRGPEKIRPECFELLRYVKGQGVGVFQVRKYTGANTQIIPMKYVIL
KRMIVRNAAKDTATKAERINILEEVKHPFIVGQIVPOTGGKLYLILEVSGGLPQMO
LEBEGIFMEDTACFYILAELISMALGHLHOKETIYRLDKPENIMLNHQSHVKLTIDGLCK
ESLHDGTVHTFGGITLEYAPALILMSRGNHRADWMSLGLADLMDLGAFFETIENKRS
KITDKILCKLIMLPYLTQEARLDILKILRNANASRIAGAPAGAGVEQAPFHEHINR
EELIARKEVEPPEKPILOESEDVQSPSKFRTCPVPSDPTSESNQVELIGTYYVA
PSLSEVSEKESPEPKIRSRPRFVGSRRPVSQVSKSGPDGFMWGAASATNAPOTPV
YPMETISGICMDVTTSGEASAPLPIQPNISGPYKCAFPPIISKRPHELRNML"

```

Query Match Similarity	2.2%	Score 24	DB 10	Length 1800
Best Local Similarity	100.0%	Pred. No. 0.36		
Matches 24	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      642 GACCTGAAGCCGGAGAATCATG 665
          |||
Db       716 GACCTGAAGCCGGAGAATCATG 739
```

[illegible]

DEFINITION
Mus musculus RIKEN cDNA 2610318i15 gene, mRNA (cDNA clone MGC:47115
IMAGE:4010710), complete cds.
ACCESSION
BC038491
CD038491 1 CT:325123AF
DEFINITION
Mus musculus RIKEN cDNA 2610318i15 gene, mRNA (cDNA clone MGC:47115
IMAGE:4010710), complete cds.
ACCESSION
BC038491
CD038491 1 CT:325123AF

ORGANISM	Mus musculus
SOURCE	Mus musculus (house mouse)
KEYWORDS	MG.
VERSION	BC036791.1
GI	43342343

REFERENCE
1 (bases 1 to 1846)
Eukaryotes; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Klausner, R.D., Collins, F.S., Wagner, L., Shamen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhac, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.

Stapleton, M., Soares, M. B., Donald, M. F., Casavant, T. L.,
Schaefer, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J.,

McKernan, K. J., Malek, J. A., Gunnarone, P. H., Richards, S.,
Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulik, S. W.,
Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A.,

Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dikren, M. C., Rodermus, R., Crawford, T., Schwartz, I., Wherry, B. M.

TITLE
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

UOJUNAD U.S.A. 99 1607-16305 (2002)
FICOR.NELL. Acad. Sci.
22388257
PUBMED 12477932
2 (bases 1 to 1846)
REFERENCE

Authors
 Title
 Direct Submission
 Submitted (04-OCT-2002)
 National Institutes of Health, Mammalian
 Gene Collection (MGC),
 Cancer Genomics Office, National Cancer

REMARK	USA
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mc@nih.gov

Ahter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stancitop, S., Thomas, P.J., Touchman, J.W., Tsurugon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRAC Plate: 80 Row: 1 Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.

FEATURES

source
1. 1846
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech 11"
/db_xref="taxon:10090"
/clone="MGC:47115 IMAGE:4010710"
/tissue_type="Mammary tumor metastasized to lung."
MMTV-LTR/mtl model. Expression driven by an MMTV-LTR enhancer.
/clone_1db="NCI CGAP_Ln30"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
1. 1846
/gene="2610318115Rk1k"
/db_xref="locusid:72508"
/db_xref="WGI:1919758"
75. 1652
/product="2610318115Rk1k protein"
/protein_id="AAH38491.1"
/db_xref="GI:23512346"
/db_xref="locusid:72508"
/translation="MRRRRRRDGFYLPAPRRREADMAGVFDIDIDPEDNAGEDEL
EEGQNLSDHGVGVPYELGMECHKEFEISVNRGEKIRPECFELLRLVKGKY
GKFOVKVYGTANTGKI FAKVLKKAMIVENAKDTATKERNITLSEVAFIVDLY
AFOTGKLYLILEYLSGSELFMOLERGIEMEDTACFYLAESMALGHOGKIYRD
LKPENIMLNHOGHVKLTDGELCKESIHDGVTHTPGCTIENVAPEILMSGNRAVDW
WSLGALMYDMITGAPPTGPNRKKTIKIKCKNLPPYTOEARDLKLLKRNAS
RLGAGPDAGEVQAHPPFRHINWELLARKVPEPRFLQSEEDVSGPDSKTRQTPV
DSPDDSTLSEANQVFLGTYVAPSVLESVKEKSPKIRSRFRIGSPRTVSPVK
FSPDFWGRGASSTANPQTPVEYPMETSGIEQMDVTSGEASAPLPPIRQNSGPYK
QAFPMISKREPHRLNL"
345. 1130
/note="S TKs; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CCD:smart00220"

ORIGIN

Query Match 2.2%; Score 24; DB 10; Length 1846;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGGAGAACATCATG 665
|||||
Db 726 GACCTGAAGCCGGAGAACATCATG 749
|||||
RESULT 43
LOCUS RAT56PK 2287 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat S6 protein kinase mRNA, complete cds.
ACCESSION M58340 M37777

VERSION M58340.1 GI:206841
KEYWORDS S6 protein kinase, insulin/mitogen stimulated protein kinase; p70

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus
1 (bases 1 to 2287)

Banerjee, P., Ahmad, M.F., Grove, J.R., Kozlosky, C., Price, D.J. and
Avruch, J.

Molecular structure of a major insulin/mitogen-activated 70-kDa S6
protein kinase

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (21), 8550-8554 (1990)

MEDLINE 91046033

COMMENT 2236064

FEATURES
source
1. 2287
/location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/cell_line="hepatoma line H4IIEC3"
/tissue_1db="lamda g11 pb4"
22. 1599
/codon_start=1
/product="S6 protein kinase"
/protein_id="AA42104.1"
/db_xref="GI:206842"

CDS

1. 2287
/translation="MRRRRRRDGFYLPAPRRREADMAGVFDIDIDPEDNAGEDEL
EEGQNLSDHGVGVPYELGMECHKEFEISVNRGEKIRPECFELLRLVKGKY
GKFOVKVYGTANTGKI FAKVLKKAMIVENAKDTATKERNITLSEVAFIVDLY
AFOTGKLYLILEYLSGSELFMOLERGIEMEDTACFYLAESMALGHOGKIYRD
LKPENIMLNHOGHVKLTDGELCKESIHDGVTHTPGCTIENVAPEILMSGNRAVDW
WSLGALMYDMITGAPPTGPNRKKTIKIKCKNLPPYTOEARDLKLLKRNAS
RLGAGPDAGEVQAHPPFRHINWELLARKVPEPRFLQSEEDVSGPDSKTRQTPV
DSPDDSTLSEANQVFLGTYVAPSVLESVKEKSPKIRSRFRIGSPRTVSPVK
FSPDFWGRGASSTANPQTPVEYPMETSGIEQMDVTSGEASAPLPPIRQNSGPYK
QAFPMISKREPHRLNL"

ORIGIN

Query Match 2.2%; Score 24; DB 10; Length 2287;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGGAGAACATCATG 665
|||||
Db 673 GACCTGAAGCCGGAGAACATCATG 696
|||||

RESULT 44
AL604063 146759 bp DNA linear ROD 29-NOV-2001

LOCUS Mouse DNA sequence from clone RP23-467012 on chromosome 11,
complete sequence.

DEFINITION AL604063

ACCESSION AL604063

VERSION AL604063.4 GI:17221230

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 146759)

LOVELL, J.

Direct Submission

Submitted (29-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk

On Nov 30, 2001 this sequence version replaced gi:1644787.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission

COMMENT

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep R23-467J12 is constructed by the group of Pieret de Jong. For further details see <http://www.chori.org/bagpac/home.htm> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP23-467J12. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true left end of clone RP23-366M19 is at 144760 in this sequence. The true right end of clone RP23-259D24 is at 2000 in this sequence.

FEATURES	Location/Qualifiers
source	1. .146759

ORIGIN

Query Match	2.2%	Score 24	DB 10	Length 146759
Best Local Similarity	100.0%	Pred. No. 0.25		
Matches 24	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	642	GACCTGAAGCCGGAGACATCATG	665
Db	35687	GACCTGAAGCCGGAGACATCATG	35664

RESULT 45	AC128095	AC128095	243780 bp	DNA	linear	HTG 13-NOV-2007
LOCUS	AC128095					
DEFINITION	Rattus norvegicus clone CH230-319P22, *** SEQUENCING IN PROGRESS					
ACCESSION	AC128095					
VERSION	AC128095.3					
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.					
SOURCE	Rattus norvegicus					
ORGANISM	Rattus norvegicus (Norway rat)					

REFERENCE
AUTHORS

Muzny, D. Walter, Weizker, M. Lee, Adimzon, S., Adams, C., Ader, J.,
 Allen, C., Allen, H., Alsbrook, S., Amin, A., Anguiano, D.,
 Arayalespech, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H.,
 Baldwin, B., Bandaranaike, D., Barber, M., Barnesread, M., Bernhead,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Bunay, C., Burch, P., Burrill, K.,
 Cardenas, V., Carter, K., Cavarros, I., Cesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, X., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedetch, D.,
 Delgado, O., Denison, S., Bertram, C., Ding, Y., Dinh, H., Diya, K.,
 Draper, H., Dugan-Socha, S., Dunn, A., Dublin, K., Dural, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabriel, A., Gante, R., Garcia, A., Garner, T., Garcia, W.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guerra, W.,

JOURNAL TITLE	Direct Submission
JOURNAL REFERENCE	Unpublished 2 (bases 1 to 243780)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUL-2002) Human Genome Sequencing Center, Department

REFERENCE 3 (bases 1 to 243780)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:22195451.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GZOP
Center clone name: CH230-319P22
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329
Consensus quality: 185810 bases at least Q40
Consensus quality: 190701 bases at least Q30
Consensus quality: 194362 bases at least Q20

Estimated insert size: 192713; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 134122: contig of 134122 bp in length
* 134123 134222: gap of unknown length
* 134223 223289: contig of 89067 bp in length
* 223290 223388: gap of unknown length
* 223390 243780: contig of 20391 bp in length.

FEATURES

SOURCE

1. .243780
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-319P22"

misc_feature

1. .1586
/note="wgs_end_extension"
clone_end:Sp5"
7959. .8865
/note="clone boundary"
clone_end:Sp6"

misc_feature

50232. .51649
/note="wgs contig"
87231. .88651
/note="wgs contig"
167979. .2170229
/note="wgs contig"
191898. .193696
/note="wgs contig"

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match

Best Local Similarity 2.2%; Score 24; DB 2; Length 243780;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

642 GACCTGAAGCCGGAACAATCATG 665
|||||

DB

174521 GACCTGAAGCCGGAACAATCATG 174544
|||||

Search completed: June 30, 2004, 20:19:56
Job time : 4628 secs

Thu Jul 1 09:50:14 2004

figure9.rnt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 01:24:35 ; Search time 178 Seconds
(without alignments)
1125.480 Million cell updates/sec

Title: FIGURE9
Sequence: 1826
1 XXVIGLIGRSSPXQLDSLS.....GXEMSLGLDMPGHQXRTX 361

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-O=/cgm2.1/USPTO.spool/GEBSREYSUS41/rnat.29062004.143151.24577/app.query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastp -SUFFIX=frn -MINMATCH=0 -1-LOOPEXT=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=GEBSREYSUS41@cgm.1.1.56 @rnat.29062004.143151.24577 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONELIG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Issued Patents NA:*

1: /cgm2_6/prodata/2/ina/5A_COMB.seq:*

2: /cgm2_6/prodata/2/ina/5B_COMB.seq:*

3: /cgm2_6/prodata/2/ina/6A_COMB.seq:*

4: /cgm2_6/prodata/2/ina/6B_COMB.seq:*

5: /cgm2_6/prodata/2/ina/POCUS_COMB.seq:*

6: /cgm2_6/prodata/2/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	722	39.5	1429	2	US-09-159-385-4
2	722	39.5	1429	2	US-09-186-277-4
3	632	34.6	2132	2	US-09-159-385-3
4	632	34.6	2132	2	US-09-186-277-3
5	428.5	23.5	4935	2	US-08-631-097-3
6	428.5	23.5	5288	3	US-08-810-712-9
7	235.5	12.9	5288	4	US-09-428-711A-15
8	222	12.2	2061	4	US-09-800-960-1
9	222	12.2	2061	4	US-10-096-960-1
10	218.5	12.0	1074	4	US-09-733-388-3
11	218.5	12.0	1158	4	US-09-733-388-1
12	218.5	12.0	1671	4	US-09-733-388-5

13	218.5	12.0	1733	4	US-09-620-312D-526	Sequence 526, App
14	215	11.8	2908	4	US-09-930-181-1	Sequence 1, Appl
15	214.5	11.7	1694	4	US-09-579-664B-3	Sequence 3, Appl
16	211.5	11.6	1282	4	US-08-878-989-12	Sequence 12, Appl
17	211.5	11.6	1282	3	US-09-272-796-12	Sequence 12, Appl
18	211.5	11.6	1282	4	US-09-016-434-953	Sequence 953, App
19	211	11.6	781	4	US-09-016-434-851	Sequence 26, Appl
20	205.5	11.3	4162	4	US-08-459-448A-26	Sequence 26, Appl
21	205.5	11.3	4162	3	US-08-459-559A-26	Sequence 26, Appl
22	205.5	11.3	4162	3	US-08-459-504B-26	Sequence 26, Appl
23	205.5	11.3	4162	3	US-08-459-444-26	Sequence 26, Appl
24	205.5	11.3	4162	3	US-08-547-422-26	Sequence 26, Appl
25	203.5	11.1	4165	1	US-07-951-715A-26	Sequence 26, Appl
26	203.5	11.1	3364	4	US-09-930-181-3	Sequence 3, Appl
27	200.5	11.0	1480	4	US-09-016-434-1454	Sequence 1454, Ap
28	199.5	10.9	5708	4	US-09-566-921-21	Sequence 21, Appl
29	197.5	10.6	1732	4	US-09-430-564-1	Sequence 1, Appl
30	191.5	10.5	3472	2	US-08-715-568A-2	Sequence 2, Appl
31	191.5	10.5	8906	2	US-08-826-267-1	Sequence 1, Appl
32	190.5	10.4	2454	4	US-09-992-481-3	Sequence 3, Appl
33	190.5	10.4	2824	4	US-09-992-481-5	Sequence 5, Appl
34	189	10.4	2175	4	US-09-984-890-1	Sequence 1, Appl
35	188.5	10.3	5207	4	US-09-858-664A-1	Sequence 1, Appl
36	188.5	10.3	5207	4	US-10-274-978-1	Sequence 1, Appl
37	188.5	10.3	5207	4	US-10-274-978-3	Sequence 3, Appl
38	187.5	10.3	2165	4	US-09-620-312D-809	Sequence 809, App
39	185.5	10.2	2673	4	US-09-428-711A-1	Sequence 1, Appl
40	185.5	10.2	2902	4	US-09-579-664B-4	Sequence 4, Appl
41	183.5	10.0	3995	4	US-09-428-711A-13	Sequence 13, Appl
42	181	9.9	1637	2	US-08-966-316-10	Sequence 10, Appl
43	178	9.7	2298	4	US-09-975-326-3	Sequence 3, Appl
44	178	9.7	2298	4	US-10-217-357-3	Sequence 3, Appl
45	177.5	9.7	2301	4	US-09-975-326-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-159-385-4
; Sequence 4, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAT, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(1353)
US-09-159-385-4

Alignment Scores:

Pred. No.: 1,01e-52
Score: 722.00
Percent Similarity: 73.82%
Best Local Similarity: 59.96%
Query Match: 39.54%
DB: 2
Gaps: 5

FIGURE9 (1-361) x US-09-159-385-4 (1-1429)

QY 99 PheAlaIleValArgLeuGys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
|||||

```

Db      79  TTGGCATTGTCGCAAGTGCACAGAGCGGACGAGCATGAGTATGACCAAGTTC 138
Qy      117  ILeIySLySArGAgLpProSerSerArgGlyValSerArgGluGluIleGlyArg 136
Db      139  ATCAAGAAAGCGCGCCCTGCCATCCAGCCGCGCGGTGTGAGCGGAGGAGATCAAGC 198
Qy      137  GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db      199  GAGGTGAGCATCCCTGCCGAGATCCGCCACCCCAACATCAACATGACATGACGTTTC 258
Qy      157  GluAsnLySThrAspValValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db      259  GAGAACAAAGACAGATGTGTGTGATCTGAGAGTGTGTGCGGTGCGAGTTCGAC 318
Qy      177  PheLeuAlaGlyLysAsp-His***GlnArgMetArgProArgSerSerSerArgSe 196
Db      319  TTCCTGGCCGAGAGAGATCATTTGACGAGAGATGAGCCAGAGTTCCTCAACAATC 378
Qy      196  rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db      379  CTAGACGGTGTCCACTCTGCACTCCAGCGCATCCGACACTTTGACTGAGGCCGAG 438
Qy      216  gThrSerCysSerTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db      439  AACATCATGTTCGTCGACAAAGCAAGCAGCCAGCCCGCATTAAGCTCATGACTTTGG 498
Qy      236  aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSer 256
Db      499  ATCCGACAGAGATCGAGGCTGGCAGAGATTCAGAAACATCTTTGGCA----- 547
Qy      256  rSerValArgGlyArgCysGlyHisIlePro-IleGly***IleLeuHisGlyLeuGlyL 276
Db      548  -----CACCCGAGT----- 557
Qy      276  euthrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThr 296
Db      558  -----TGT-----CGCCCGAGATCGTGAACATGATGAGCCACTT 591
Qy      296  rPLeuGlySer***TyValGluHisArgArgHisHisLeuHisProValSerAla***A 316
Db      592  GGCTTGAGGCTGACATGTGAGCATGCGCTCATCACTACATCCCTCTGAG----- 644
Qy      316  spGlyGlnGlyProGlnThrValProAlaArgGly 327
Db      645  -----CGAGCGTCCCATTCCTGGG 665

RESULT 2
US-09-186-277-4
; Sequence 4, Application US/09186277
; Patent No. 6171841
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186,277
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(1353)
US-09-186-277-4

Alignment Scores:
Pred. No.: 1,01e-52 Length: 1429
Score: 722.00 Matches: 163

```

```

Percent Similarity: 73.82% Conservative: 9
Best Local Similarity: 69.96% Mismatches: 21
Query Match: 39.54% Indels: 40
Db: 3 Gaps: 5

FIGURE9 (1-361) x US-09-186-277-4 (1-1429)
Qy      99  PheAlaIleValArgLysCys-----LysGlyThrArgIleMetGlyTrpAlaAlaLysPhe 116
Db      79  TTGGCATTGTCGCAAGTGCACAGAGCGGACGAGCATGAGTATGACCAAGTTC 138
Qy      117  ILeIySLySArGAgLpProSerSerArgGlyValSerArgGluGluIleGlyArg 136
Db      139  ATCAAGAAAGCGCGCCCTGCCATCCAGCCGCGCGGTGTGAGCGGAGGAGATCAAGC 198
Qy      137  GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db      199  GAGGTGAGCATCCCTGCCGAGATCCGCCACCCCAACATCAACATGACATGACGTTTC 258
Qy      157  GluAsnLySThrAspValValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db      259  GAGAACAAAGACAGATGTGTGTGATCTGAGAGTGTGTGCGGTGCGAGTTCGAC 318
Qy      177  PheLeuAlaGlyLysAsp-His***GlnArgMetArgProArgSerSerSerArgSe 196
Db      319  TTCCTGGCCGAGAGAGATCATTTGACGAGAGATGAGCCAGAGTTCCTCAACAATC 378
Qy      196  rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db      379  CTAGACGGTGTCCACTCTGCACTCCAGCGCATCCGACACTTTGACTGAGGCCGAG 438
Qy      216  gThrSerCysSerTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db      439  AACATCATGTTCGTCGACAAAGCAAGCAGCCCGCATTAAGCTCATGACTTTGGC 498
Qy      236  aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSer 256
Db      499  ATCCGACAGAGATCGAGGCTGGCAGAGATTCAGAAACATCTTTGGCA----- 547
Qy      256  rSerValArgGlyArgCysGlyHisIlePro-IleGly***IleLeuHisGlyLeuGlyL 276
Db      548  -----CACCCGAGT----- 557
Qy      276  euthrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThr 296
Db      558  -----TGT-----CGCCCGAGATCGTGAACATGATGAGCCACTT 591
Qy      296  rPLeuGlySer***TyValGluHisArgArgHisHisLeuHisProValSerAla***A 316
Db      592  GGCTTGAGGCTGACATGTGAGCATGCGCTCATCACTACATCCCTCTGAG----- 644
Qy      316  spGlyGlnGlyProGlnThrValProAlaArgGly 327
Db      645  -----CGAGCGTCCCATTCCTGGG 665

RESULT 3
US-09-159-385-3
; Sequence 3, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT APPLICATION NUMBER: US/09/159,385
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA

```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)..(1455)
US-09-159-385-3

Alignment Scores:
Pred. No.: 9,57e-45 Length: 2132
Score: 632.00 Matches: 155
Percent Similarity: 62.69% Conservative: 13
Best Local Similarity: 57.84% Mismatches: 46
Query Match: 34.61% Indels: 54
DB: 2 Gaps: 9

FIGURE9 (1-361) x US-09-159-385-3 (1-2132)

QY 99 PheAlaIleValArgIysCys-----LysGlyThrGlyMetGluTyrAlaAlaIysPhe 116
Db 163 TTTCGATCGTGGGAAAGTCCGGCAGAAAGGCGACAGAGATACCGACCAAGTTC 222
QY 117 IleValysArgIysLeuProSerSerArgIysValSerArgGluIleGluArg 136
Db 223 ATCAAGAGAGCGCGCTGTCATCCAGCGCGGTGGGTGAGCGGAGAGATCCAGCGG 282
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 283 GAGGTAAACATCTCGCGGAGATCCGGCACCCCAACATCATCCTTGACGACATCTTC 342
QY 157 GluAsnIleThrAspValValIleuIleLeuGluLeuValSerGlyIleuPheAsp 176
Db 343 GAGAACAGAGCGAGCGTCTCTCATCTCGAGCTGTCTCGCGGAGCTCTTGAC 402
QY 177 PheLeuAlaGluIysAspHis***GlnArgMetArgProArgSerSerSerArgSe 196
Db 403 TTCTCGCGGAGAAAGTCCGTGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db 463 CTGAGCGGCTTCACTCACTGCACTCTACAGCGCATGACACATTTTACCTGAAAGCGGAA 522
QY 216 gThrSerCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 523 AACATCATGCTGCTGACAGAGACGTGCCAACCCCAACATCAAGCTTCATCGACTTCGCGC 582
QY 236 aserArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 583 ATGCGGACAGAGATCGAGCGCGGAGACGAGTTCAGAGACA-----TC 624
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
Db 625 TTTCGGACCCCGGAGCTTGTGGC----- 647
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro***AspCysIysLeu***ThrThr 296
Db 648 -----CCACAGATTGTGAATGATGAGCGCGCTG 675
QY 296 rPheGlySer***TyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316
Db 676 GCGCTGAGCGGACGACATGAGCATGCGTGTCTATCCCTATATCTC----- 722
QY 316 spgIyngIyProGln-----ThrValProAlaArgIyProGlyIleArgAlaG 333
Db 723 -----CCAGACCGGTGATCCCGTCTCTCGCGGAGAC-----CAAGCAG 762
QY 333 lYthrSerAlaAsnCysLeuHisTrpGly***GluMetSerLeu-----GlyThrLeuAl 351
Db 763 GAGACGCTCAC-----CAACATCTCAGCCGTAACTACATTCGACGAGAGTACTTC 816
QY 351 spMetProGlyProHisGln 357
Db 817 AGCAAA-----CACCAAG 827
RESULT 4

US-09-186-277-3
Sequence 3, Application US/09186277
Patent No. 6,178,41
GENERAL INFORMATION:
APPLICANT: AKIRA, SHIZUO
APPLICANT: KAMAI, TARO
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
FILE REFERENCE: 081356/0128
CURRENT APPLICATION NUMBER: US/09/186, 277
CURRENT FILING DATE: 1998-11-05
EARLIER APPLICATION NUMBER: JP97/261589
EARLIER FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2132
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)..(1455)
US-09-186-277-3

Alignment Scores:

Pred. No.: 9,57e-45 Length: 2132
Score: 632.00 Matches: 155
Percent Similarity: 62.69% Conservative: 13
Best Local Similarity: 57.84% Mismatches: 46
Query Match: 34.61% Indels: 54
DB: 3 Gaps: 9

FIGURE9 (1-361) x US-09-186-277-3 (1-2132)

QY 99 PheAlaIleValArgIysCys-----LysGlyThrGlyMetGluTyrAlaAlaIysPhe 116
Db 163 TTTCGATCGTGGGAAAGTCCGGCAGAAAGGCGACAGAGATACCGACCAAGTTC 222
QY 117 IleValysArgIysLeuProSerSerArgIysValSerArgGluIleGluArg 136
Db 223 ATCAAGAGAGCGCGCTGTCATCCAGCGCGGTGGGTGAGCGGAGAGATCCAGCGG 282
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 283 GAGGTAAACATCTCGCGGAGATCCGGCACCCCAACATCATCCTTGACGACATCTTC 342
QY 157 GluAsnIleThrAspValValIleuIleLeuGluLeuValSerGlyIleuPheAsp 176
Db 343 GAGAACAGAGCGAGCGTCTCTCATCTCGAGCTGTCTCGCGGAGCTCTTGAC 402
QY 177 PheLeuAlaGluIysAspHis***GlnArgMetArgProArgSerSerSerArgSe 196
Db 403 TTCTCGCGGAGAAAGTCCGTGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db 463 CTGAGCGGCTTCACTCACTGCACTCTACAGCGCATGACACATTTTACCTGAAAGCGGAA 522
QY 216 gThrSerCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 523 AACATCATGCTGCTGACAGAGACGTGCCAACCCCAACATCAAGCTTCATGACTTCGCGC 582
QY 236 aserArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 583 ATGCGGACAGAGATCGAGCGCGGAGACGAGTTCAGAGACA-----TC 624
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
Db 625 TTTCGGACCCCGGAGCTTGTGGC----- 647
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro***AspCysIysLeu***ThrThr 296
Db 648 -----CCACAGATTGTGAATGATGAGCGCGCTG 675

Qy 296 rpleuglyser***TyValGluHisArgArghHisLeuHisProValSerAla***A 316
Db 676 GGCTGGAGGGCGGACATGTGAGGACATCGGTCTATCTACCTTATCTCT----- 722
Qy 316 spgIyngIyProGln-----ThyValProAlaArgGlyProGlyIleArgAlaG 333
Db 723 -----CCTGACGCGNGCATCCCCGCTCTCTGGGCGACAG-----CAAGCAG 762
Qy 333 lYthrsrAlaAsnCysLysHisTrrpGly***GluMetSerLeu-----GlyThrLeuA 351
Db 763 GAGCGGTCTCAC-----CAACATCTCAGCCGTGAATACGACTTCGACGAGGAGTACTTC 816
Qy 351 spMetProGlyProHisGln 357
Db 817 AGCAA-----CACCCAG 827

RESULT 5
US-08-631-097-3
; Sequence 3, Application US/08631097
; Patent No. 5968816
; GENERAL INFORMATION:
; APPLICANT: Kimchi, Adi
; TITLE OF INVENTION: Tumor Suppressor Genes,
; TITLE OF INVENTION: Protein Encoded Thereby, and Use of Said Genes and Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wigman, Cohen, Leitner, & Myers, P.C.
; STREET: 900 17th Street, N.W., Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/631,097
; FILING DATE: 12-Apr-96
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11598
; FILING DATE: 12-Oct-94
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Herbert
; REGISTRATION NUMBER: 25,109
; REFERENCE/DOCKET NUMBER: 0744.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)463-7700
; TELEFAX: (202)473-6915
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHEICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: No. 5968816 applicable
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; STRAIN: not applicable
; INDIVIDUAL ISOLATE: not applicable
; DEVELOPMENTAL STAGE: not applicable
; TISSUE TYPE: blood
; CELL TYPE: Leucocyte
; CELL LINE: HeLa
; ORGANELLE: not applicable
; IMMEDIATE SOURCE:
; LIBRARY: not applicable

CLONE: not applicable
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: not applicable
; MAP POSITION: not applicable
; UNITS: not applicable
; FEATURE:
; NAME/KEY: Seq. ID. NO.: 3 is
; LOCATION: not available
; IDENTIFICATION METHOD: in specification
; OTHER INFORMATION: prevention of IFN-2
; OTHER INFORMATION: promoted cell death
; PUBLICATION INFORMATION: not available
; US-08-631-097-3

Alignment Scores:
Pred. No.: 9,37e-27 Length: 4935
Score: 428.50 Matches: 102
Percent Similarity: 61.1% Conservative: 30
Best Local Similarity: 47.22% Mismatches: 45
Query Match: 23.47% Indels: 41
Gaps: 2 4

FIGURES (1-361) x US-08-631-097-3 (1-4935)

Qy 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluYrAlaAlaLysPhe 116
Db 406 TTGGCGTGTGGAGAAATGCGGTGAGAAAGTACCGGCTCCGCAATTC 465
Qy 117 lIelysLysArgGluProSerSerArgArgLysLysSerArgGluGluIleGluArg 136
Db 466 ATCAGAAAGAGAGACTTAATCCAGCCGGGGGTGTGAGCCCGAGAGATCGACCGG 525
Qy 137 GluValSerIleLeuArgGluIleArgHisProAsnIleThrLeuHisAspValPhe 156
Db 526 GAGGTGACATCCCTGGAAGAGATCCAGACCCCAATCTATCACCCTCGACGAGTCTAT 585
Qy 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 586 GAGAAACAAGAGGAGCATCTGATCTGATCTGGAATCCGTCGAGTGGCGAGCTGTTGAC 645
Qy 177 PheLeuIleGluLysAspHis***Gln-ArgMetArgProArgSerSerSerArgSe 196
Db 646 TTCTAGCTGAAGAGATCTTAACTGAAGAGACCAATGATTTCTCAACAAATT 705
Qy 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 706 CTATATGCTGTTAATCTACCTGCACTCCCTCAATGCGCCCACTTATCTTAAGCTGAG 765
Qy 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236
Db 766 AACATATGCTTTGGATAGAAATGTCCTCA----- 796
Qy 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 797 ----AACTCGGATCATGATCATG-----ACTTTGGA-AATGATTTAA 836
Qy 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyLe 276
Db 837 AAACATATTTGGG----- 849
Qy 276 uThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrTr 296
Db 850 -----ACTCGAGATTTCGCTCTCTGGA-GAATGCAATTAATGAACCTCTTG 895
Qy 296 rpleuglyser***TyValGluHisArgArghHisLeuHisPro 311
Db 896 GTCTTGAGGACAGATATGTGAGTATCGGGGTAAATTAACCTATATCCT 941

RESULT 6
US-08-810-712-9
; Sequence 9, Application US/088107129

```
; Patent No. 6160106
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. LTD
; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
; TITLE OF INVENTION: Use of said Genes and Proteins
; FILE REFERENCE: sequence list
; CURRENT APPLICATION NUMBER: US/08/810,712G
; EARLIER FILING DATE: 1997-03-03
; EARLIER APPLICATION NUMBER: PCT/US94/11598
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 5886
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (337)..(4605)
; US-08-810-712-9

Alignment Scores:
Pred. No.: 1,21e-26 Length: 5886
Score: 428.50 Matches: 102
Percent Similarity: 61.11% Conservative: 30
Best Local Similarity: 47.22% Mismatches: 45
Query Match: 23.47% Indels: 41
DB: Gaps: 4

FIGURE9 (1-361) x US-08-810-712-9 (1-5886)
Qy 99 PhenAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 406 TTTCGGGTGTGAAGAAATGCCGAGAAAGTACCGGCTCCAGTATCCCGCAATTC 465
Qy 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
Db 466 ATCAAGAAAGAGAGACTAATGTCACGGCGGGGGTGTAGCGCGACAGCATCGACGG 525
Qy 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 526 GAGGTGCGATCTCTGAAGAGATCCAGCACCCCAATGTCATCACCTCGACGAGCTCTAT 585
Qy 157 GluAsnLysThrAspValValIleuIleLeuGluLeuValSerGlyGlyLeuPheArg 176
Db 586 GAGAACAGACGACGCTCATCTGAACTCTTCAGCTGCGGCGAGCTGTTGAC 645
Qy 177 PheLeuAlaGluLysAspHis***Gln-ArgMetArgProArgSerSerSerArgSse 196
Db 646 TTCTTACCTGAAGAAATCTTTAAGTGAAGAGAACAACTGAATTTCTCAACAATTC 705
Qy 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 706 CTTAATGAGTGTACTACCTGCACTCCCTTCAATCGCCCACTTGTATCTTAAGCTGAG 765
Qy 216 gTThrSerCysStrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 766 AACCTAATGCTTTGGATGAAATGTCCTCA----- 796
Qy 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgLysSerSe 256
Db 797 -----AACCTGGATCAAGATCATG-----ACCTTGA-AAATGAATTTAA 836
Qy 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyLe 276
Db 837 AAACATATTGGG----- 849
Qy 276 uThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrTr 296
Db 850 -----ACTCCAGAGTTTGGCCCTCCGGA-GATAGTCAAATATGAACCTCTTG 895
Qy 296 pLeuGlySer***TyrValGluHisArgArgHisHisLeuHisPro 311
Db 896 GTCTTGAGGACAGATATGTCAGTATCGGGGGTAAATACCTATATCTCT 941
```

```
Db 896 GTCTTGAGGACAGATATGTCAGTATCGGGGGTAAATACCTATATCTCT 941
RESULT 7
US-09-428-711A-15
; Sequence 15, Application US/09428711A
; Patent No. 6358720
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masaaki
; APPLICANT: Shiraesawa, Takaji
; APPLICANT: Tokumitsu, Hiroshi
; APPLICANT: Teruhisa
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
; FILE REFERENCE: 06501-045001
; CURRENT APPLICATION NUMBER: US/09/428,711A
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: PCT/JP96/01246
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: JP 9/124798
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 5228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (269)...(3418)
; US-09-428-711A-15

Alignment Scores:
Pred. No.: 3.65e-10 Length: 5228
Score: 235.50 Matches: 99
Percent Similarity: 39.70% Conservative: 32
Best Local Similarity: 30.00% Mismatches: 140
Query Match: 12.90% Indels: 61
DB: Gaps: 11

FIGURE9 (1-361) x US-09-428-711A-15 (1-5228)
Qy 39 GlyProArgCysArgAlaSerValLeuSerLeu-----LysAla 51
Db 194 GGGCCGGCCCTCCGCTAGTCCCGCGGCTTGCCCGCCGACCCCGCCCGCGCCCGCC 253
Qy 52 IleProLysCysLeuSer***GlySerLeuGlySerSerValValGlyAsnLeuGly 71
Db 254 CGGCCCGGCTGGCCATGAGCGCCGCGCGGACAGACCGTGGGCAAGTTGAG 313
Qy 72 ThrAspAlaGlnArgLeuGlyTyrArgValLeuProSerSerGlySerAlaLeuSer 91
Db 314 TTCTCCCGCAGAGACTG----- 331
Qy 92 CysSerPheProHisSerGlyPheAlaIleVal-----ArgLysCysLysGlyThr 108
Db 332 -----ATCGGCCAGCGCGCCTTCGCGTGGCTTCAAGGCCCGCCAGAGAGAC 385
Qy 109 GlyMetGluTyrAlaAlaLysPheIleLysArgArgLeuProSerSerArgArgGly 128
Db 386 GATTGGAGGTGCGCTGCAAGTGCATTAACAAGAACTCGGCAAG----- 433
Qy 129 ValSerArgGluGluIleGluArgGluValSerIleLeuArgGluIleArgHisProAsn 148
Db 434 ---TCTCAGAGCGCTCGGGAGAGAAATCAAAATCTCGAAGAACTGAACATGAATAAC 490
Qy 149 IleIleThrLeuHisAspValPheGluAsnLysThrAspValValLeuIleLeuGluLeu 168
Db 491 ATCGTGGCCCTGTAGACTTCCAGAAATGCTAATTTGCTTACCTGTTATGAGTAC 550
Qy 169 ValSerGlyGlyGluLeuPheAspPheLeuAlaGluLysAspHis***GlnArgMetArg 188
Db 551 TGCACAGGTGGGAGACTCGCGACACTACCT-GCAGCCCAATGGGACGCTGAGGAGAG 608
Qy 189 ProArgSerSerSer---SerArgSerTrpThrValSerThrCysThrProSerAla 207
```

```

Db      609 CCATCAGAGCTCTTCTCTCAGACAGATCCGGGCGCCATGCGCTTCTGCACAGCAAGAGCA 668
Qy      208 SerArgThrLeu***SerArgArgThrSerCysThrPthSerMetGlnProAla 227
Db      669 TCATCCACCGGAGCTTGAACCGCAACATCTCTGCTCCAAAC-----CGCGCGCC 722
Qy      228 HisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerArgProValAlaSerSer 247
Db      723 GCCCGCCCAACCCCAACAGCATCCGCGCTCAAGA-----TCG 758
Qy      248 ArgThrSerLeuAlaArgGlnSerSerSerValArgGly-ArgCysGlyHisHisProI 267
Db      759 CTGACTTCGGCTTGGCGCGTACTCC-----AGACCAACATGATGCGCGCAC----- 807
Qy      267 eGly**IleLeuHisGlyLeuGlyLeuThrCysLeuAsnAsnProValPheHisSerP 287
Db      808 -----ACTCTCGGCTCCCGCATATCATGCGCC 836
Qy      287 C***AspCysLeuLeu**ThrThrTrpLeuGlySer**Ty-ValGlnHisArgArgH 307
Db      837 CCGAGGTCACTATCTCCAGCATACGACGGAAGCGGACCTGTGAGCATCGGACCA 896
Qy      307 SH:LeuHisProValSerAla**AspGlyGlnGlyPro---GlnThrValProAlaAr 326
Db      897 TCGTCTA---CGAGTGCCTACGCGGGAAGCGCCCTTCCAGCCAGAGCCCGCAGACC 953
Qy      326 gGlyProGlyTrpLeuArgAlaGlyThrSerAlaAsnCysLysHisTrpGly**GluMetSe 346
Db      954 TGCCCTGCTTCTACGAAAGAACAGACGTGTGTCGCCACATCCCGGAGAGACCTCG 1013
Qy      346 rLeuGlyThrLeuAspMetProGlyPro 355
Db      1014 CCGCGCTGCGGACACTCTCTGCGCCT 1041

RESULT 8
US-09-800-960-1
; Sequence 1, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Human
US-09-800-960-1

Alignment Scores:
Pred. No.: 1,34e-09 Length: 2061
Score: 222.00 Matches: 83
Percent Similarity: 44.00% Conservative: 38
Best Local Similarity: 30.18% Mismatches: 93
Query Match: 12.16% Indels: 63
DB: 4 Gaps: 11

FIGURE9 (1-361) x US-09-800-960-1 (1-2061)
Qy      61 LeuGlySerSerValValAlaGlyAsnLeuGlyThrAspAlaGlnArg----- 76
Db      23 GTCTCCCTCTCTCTCTCTCGCGCGCGCGGTACTGTGACCGACGCTCGCGCGG 82
Qy      77 LeuGlyTrpArgValLeu---ProSerSerGlySerAlaAlaLeuSerCysSer----- 93
Db      83 CTGACCGCGCGCGTCTCGCGCGCGCGCATGAGCCACCGCCACCTGACCCGCTTC 142
```

```

Qy      94 -----PheProHisSerGly-----PheAlaIleValArgLys 104
Db      143 ACCGACGACTACCACTCTTCTGAGAGAGCTTGCAAGGCTGCTCTCTGCTCGGAG 202
Qy      105 Cys-----LysGlyThrGlyMetGlyTrpAlaAlaLysPheIleLysArgArgLeu 122
Db      203 TGTGTGAAGAAACCTTCACGACGAGAGTACGAGAAATATCATCATATACCAAGAAATTG 262
Qy      123 ProSerSerArgArgGlyValSerArgGluGluIleGluArgGluValSerIleLeuArg 142
Db      263 ---TCTGCCGG-----GATCACCAAGAACTGAGACGTGAGGCTCGGATATGTGCA 310
Qy      143 GluIleArgHisProAsnIleIleThrLeuHisAspValPheGluAsnLysThrAspVal 162
Db      311 CTTCGAACATCCCAACATGCTGCTCCATACAGATTCTTGAAGAAAGGTTTCAC 370
Qy      163 ValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAspPheLeuAlaGluLysAsp 182
Db      371 TACCTCGTGTGACCTTGTACCGGCGGAGAGTGTGTAAGACATTGTGCGCAGAG 430
Qy      183 His***Gln-ArgMetArgProArgSerSerSerSerArgSerTrpThrValSerThrTh 202
Db      431 TACTACACTGAAGACAGATGCGCAGCCACTGTATACATCAGATTCTGAGAGAGTTAACAC 490
Qy      202 rCysThrProSerAlaSerArgThrLeuThr**SerArgArgThrSerCysTrpThr 222
Db      491 ATCCACGACGATGATGTCTCCACAGGACCTGAAGCTGGAAC-----TCTGCTGG-- 545
Qy      222 rSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerAr 242
Db      546 -----CGAGTAAATGCAAG 559
Qy      242 gProValAlaSerSerArgThrSerLeuAla-----ArgGlnSerSerSe 257
Db      560 GGTGCGCGCGGTCAAGACTGTGATTTTGCTTACGCCATCCAGATACAGGAGACAGACAG 619
Qy      257 rValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyLeuTh 277
Db      620 G-----CTTGTTGGT 631
Qy      277 rCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrTrpLe 297
Db      632 TTTGCTGAC-ACCCAGGATCTGTGCTCCCTGAGGCTTGAAG-AGATCCCTATGGA 689
Qy      297 uGlySer**Ty-ValGlnHisArgArgHisHisLeuHisPro 311
Db      690 AACCTGTGATATCTGCGCGCTCGCGGCTCATCTGATATCTCT 732

RESULT 9
US-10-096-960-1
; Sequence 1, Application US/10096960
; Patent No. 6664085
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001158DIV
; CURRENT APPLICATION NUMBER: US/10/096,960
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-096-960-1

Alignment Scores:
Pred. No.: 1,34e-09 Length: 2061
Score: 222.00 Matches: 83
```



```

OY 106 LysGlyThnGlyMetClnuYrAlaIleAlaLysPheIleuLYsAspArgLeuProSer 125
Db 127 AAGGCACTGGCAAGCTCTTCTGCTGAAAGTATACCTTAAGAAAGCGCTGAAGGC--- 183
OY 126 ArgArgGlyValSerArgGluIleGluArgGluValSerIleLeuArgGluIleArg 145
Db 184 -----AAGAAAGACAGCAATAGAAATGAGATGAGATGACCGCTCTCGAAGAAATTAG 231
OY 146 HisProAlaIleIleThrLeuHisAspValAPheGluAsnIlySthAspValIleuIle 165
Db 232 CATGAAATATTCTGTTCTCCCTCGAAGACATTATGAAACCCAAATACCTTACTTGTC 299
OY 166 LeuGluIleuValSerGlyGlyGluLeuPheAspPheLeuAlaGluIys-AspHis***GI 185
Db 292 ATCAGCTGGTGTCCCGGTGAAGAGCTGTTGACCGGAATGTGAGAAAGGGATTATACA 351
OY 185 nAcMetArgProArgSerSerSerSerArgSerIleThrValSerIleThrCysThrPr 205
Db 352 GAGAGAGATGCCAGCACTGTATCGGCAAGCTTGAGCGCGGTACTATCTCCACAGA 411
OY 205 oSerAlaSerArgThLeuThr***SerArgArgThrSerCysStrpThrSerMetGI 225
Db 412 ATGGCGATCTCTCCACAGACCTCAAGCCGCAAAATCTCTGTACTACAGTCAAGATGAG 477
OY 225 n---ProAlaHisAlaLeuSerSerSerThrIleuAlaSerArgIleuArgIlySerArgProVa 244
Db 472 GAGTCCAAATA-ATGATCAGTGACTTGATGTTCMAAAATGAGGGCCAAAGAGATGT 530
OY 244 AlaSerSerArgThSer-----IleuAlaArgIle 255
Db 531 GATGTCCACGCTGTGGAATCCAGGCTATGTGCTCTGTAAGTCTCGGCCGAAACC 599
OY 255 rSerSerValArgIlyArgCysGlyHisHisProIleGly***IleuHisGlyLeuGI 275
Db 591 TTACAGGAAGACCGTTGACTGC-----TGCATCATCGGAGGATGGCTAC- 633
OY 275 yLeuThrCysLeuHisAsnProValPheHisSerPro***AspCysIlyLeu***ThrTh 299
Db 637 -ATGTGCTCTGCGGCTACCTCTCTTTATGATGAATAAGTACGTCCAAAGCTTTTGAGCA 699
OY 295 rTrpLeuGlySer***Tyr-----ValGluHisArgAr 306
Db 696 GATCTCCAAAGCGGAATATAGATTGACTCTCCCTACTGGATGCATCTCGCACTGTC 755
OY 306 gHisHisLeuHisProValSerLa***AspGlyGlnGlyProGlnThrValProAlaAr 326

```

Db 756 -AAAGACTTCAT-----TCGAACCTGATGAGAACCCGAA-----TA 796
Qy 326 GGLPProGlyIleArgIleGlyThrSerAlaAsnCysLysHisTrp 341
Db 797 AAAGATACACGTGTGAGCAGGACGCTCGACCCATGATGCTGG 842

RESULT 11

US-09-733-388-1

Sequence 1, Application US/09733388

Patent No. 6602698

GENERAL INFORMATION:

APPLICANT: Donoho, Greg

APPLICANT: Scoville, John

APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Friedrich, Glenn

APPLICANT: Zambrowicz, Brian

APPLICANT: Abuin, Alejandro

APPLICANT: Sands, Arthur T.

TITLE OF INVENTION: No. 6602698e1 Human Kinase Proteins and

TITLE OF INVENTION: Polynucleotides Encoding the Same

FILE REFERENCE: LEX-0103-USA

CURRENT FILING DATE: 2000-12-07

PRIOR FILING DATE: 1999-12-07

PRIOR APPLICATION NUMBER: US 60/169,428

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1158

TYPE: DNA

ORGANISM: Homo sapiens

US-09-733-388-1

Alignment Scores:

Pred. No.:	1,146-09	Length:	1158
Score:	218.50	Matches:	81
Percent Similarity:	44.92%	Conservative:	34
Best Local Similarity:	31.64%	Mismatches:	105
Query Match:	11.97%	Indels:	38
DB:	4	Gaps:	8

FIGURE9 (1-361) x US-09-733-388-1 (1-1158)

Qy 106 LysGlyThrGlyMetGluTyrAlaAlaLysPheIleLysLysArgLysProSerSer 125
Db 127 AAGCACTGCGCAAGCTCTTCTGTAAGTATCCCTTAAGAGGCGCTGAAGGCG--- 183
Qy 126 ArgArgGlyValSerArgGluGluIleGluArgGluValSerIleLeuArgGluIleArg 145
Db 184 -----AAGCAAGCAGCATAGAGATGAGATAGACCTCTGAGAAAGATTAG 231
Qy 146 HisProAsnIleIleThrLeuHisAspValPheGluAsnLysThrAspValValLeuIle 165
Db 232 CATGAAATATGTGTGCGCTGGAAGACATTATGAAGCCCAATGACCTGATCTTGTC 291
Qy 166 LeuGluIleValSerGlyGlyGluLeuPheAspPheLeuAlaGluLys-AspHis**G1 185
Db 292 ATGACGTGTGTGTCGGGTGAGAGCTGTGACCGATGAGTAGAGAGGGGTTTATATA 351
Qy 185 ValGlyMetArgProArgSerSerSerArgSerTrpThrValSerThrCysThrPro 205
Db 352 GAGAGAGATCCGACACTGTGATCCGACAGCTTTGAGCGCCGTGATCATCTCCACAA 411
Qy 205 oSerAlaSerArgThrLeuThr**SerArgTrpSerCysCysTrpThrSerMetG1 225
Db 412 ATGGGATCTGTCCACAGAGACCTCAAGCCGAAATCTTGTGACACAGTCAAGATGAG 471
Qy 225 n---ProAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerArgProVa 244
Db 472 GAGTCCAAAATA-ATGATCAGTACTTGTGATTGTCAAAAATGAGAGGCGAAAGAGATGT 530
Qy 244 lAlaSerSerArgThrSer-----LeuAlaArgGlnSe 255

Db 531 GATGTCACTGCTGTGAGACTCCAGCTATGTGCTCTCTGAAGTCCCGCCAGAAACC 590
Qy 255 rSerSerValArgGlyArgGlyGlyHisHisProIleIle**IleLeuHisGlyLeuG1 275
Db 591 TTACAGCAAGCCGTGACTGC-----TGTTCATCGAGTGTGTTGCTTAC----- 636
Qy 275 yLeuThrCysLeuAsnProValPheHisSerPro**AspCysLysLeu**ThrTh 295
Db 637 -ATCTGTCTCCCGGCTACCTCTTTTATGATGAAATATACCTCCAGACTTTTGACCA 695
Qy 295 rTrpLeuGlySer**Tyr-----ValGluHisArgAr 306
Db 696 GATCTCAAGCGCGAATATGATTGACTCTCCCTACTGGAGTACATCTCCGACTGTC 755
Qy 306 GHisHisLeuHisProValSerAla**AspGlyGlnGlyProGlnThrValProAlaAr 326
Db 756 -AAAGACTTCAT-----TCGAACCTGATGAGAGAACCCGAA-----TA 796

RESULT 12

US-09-733-388-5

Sequence 5, Application US/09733388

Patent No. 6602698

GENERAL INFORMATION:

APPLICANT: Donoho, Greg

APPLICANT: Scoville, John

APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Friedrich, Glenn

APPLICANT: Zambrowicz, Brian

APPLICANT: Abuin, Alejandro

APPLICANT: Sands, Arthur T.

TITLE OF INVENTION: No. 6602698e1 Human Kinase Proteins and

TITLE OF INVENTION: Polynucleotides Encoding the Same

FILE REFERENCE: LEX-0103-USA

CURRENT FILING DATE: 2000-12-07

PRIOR FILING DATE: 1999-12-07

PRIOR APPLICATION NUMBER: US 60/169,428

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 1671

TYPE: DNA

ORGANISM: Homo sapiens

US-09-733-388-5

Alignment Scores:

Pred. No.:	1,966-09	Length:	1671
Score:	218.50	Matches:	81
Percent Similarity:	44.92%	Conservative:	34
Best Local Similarity:	31.64%	Mismatches:	105
Query Match:	11.97%	Indels:	38
DB:	4	Gaps:	8

FIGURE9 (1-361) x US-09-733-388-5 (1-1671)

Qy 106 LysGlyThrGlyMetGluTyrAlaAlaLysPheIleLysLysArgLysProSerSer 125
Db 293 AAGCACTGCGCAAGCTCTTGTGTAAGTATCCCTTAAGAGGCGCTGAAGGCG--- 349
Qy 126 ArgArgGlyValSerArgGluGluIleGluArgGluValSerIleLeuArgGluIleArg 145
Db 350 -----AAGCAAGCAGCATAGAGATGAGATAGACCTCTGAGAAAGATTAG 397
Qy 146 HisProAsnIleIleThrLeuHisAspValPheGluAsnLysThrAspValValLeuIle 165
Db 398 CATGAAATATGTGTGCGCTGGAAGACATTATGAAGCCCAATGACCTGATCTTGTC 457
Qy 166 LeuGluIleValSerGlyGlyGluLeuPheAspPheLeuAlaGluLys-AspHis**G1 185


```

QY      185  nArgMetArgProArgSerSerSerSerArgSerTTPThnValSerThrThrCysThrPr 205
Db      594  GAGAAAGAGTCCGCACTCTCATCCGCGCGAGTCTCGATGCCGTATATCTATCTCCACAGA 653
QY      205  oSerAlaSerArgThrLeuThr**SerArgArgThrSerCysCysTTPThrSerMetGl 225
Db      654  ATGGGCAATTGTCCACAGGAGCCTCAAGCCGAGAGATCTCT-----TATACTACAGTCAA 707
QY      225  n-----ProAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerAr 242
Db      708  GACGAGAGAGTCCAAATA-ATGATCAGTACTTTGGCTTGTGCAAAATGAGGCGCAAGG 766
QY      242  gProValAlaSerSerArgThrSer-----LeuAlaAr 253
Db      767  AGATGTGATGTCCACGCGCTCCGCGAGCCCGAGGCTATGTTGCTCCGGAAGTTCTCCGCCA 826
QY      253  gGlnSerSerSerValArgGlyArgCysGlyHisProIleGly**IleLeuHisGl 273
Db      827  GAAACCGTACAGCAAGCTGTGACTGC-----TGATCATCGGGGTGATCGCCTAT-- 878
QY      273  yLeuGlyLeuThrCysLeuHisMetProValPheHisSerPro**AspCysIleLeu** 293
Db      879  -----ATCTTGCTCTGTGCTTACCTCTTCTTTATGATGAAAAATGACTCGAAGCTGTT 931
QY      293  *ThrThrTripleGlySer**Tyr 301
Db      932  TGAACAGATCCTCAAGGCAAGATAT 956

```

Search completed: July 1, 2004, 01:56:09
 Job time : 185 secs

Thu Jul 1 09:50:16 2004

US-10-083-641a-6.rmpb

Page 10

```

; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING
; TITLE OF INVENTION: NEUROPSYCHIATRIC
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/04702 US0
; CURRENT APPLICATION NUMBER: US/10/125,635
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 157875
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-125-835-1

```

```

Query Match      1.9%; Score 21; DB 15; Length 157875;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 919 CATCGCGCTCATCCTTACAT 939
DB 142369 CATCGCGCTCATCCTTACAT 142389

```

```

RESULT 29
US-10-029-386-23553
; Sequence 23553, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23553
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR15.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA58002.1, EVALU6 3.00e-65
; OTHER INFORMATION: SWISSPROT HIT: P53355, EVALU6 5.00e-18
; OTHER INFORMATION: NT HIT: g114670382, EVALU6 8.00e-77
; US-10-029-386-23553

```

```

Query Match      1.8%; Score 20; DB 15; Length 147;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 382 AGCCGCGCGGTGTGAGCCG 401
DB 95 AGCCGCGCGGTGTGAGCCG 114

```

```

RESULT 30
US-09-864-761-11330
; Sequence 11330, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.

```

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aewmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11330
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003669.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
; US-09-864-761-11330

```

```

Query Match      1.8%; Score 20; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 500 TGATCTTGAGAGCTGTGCTCC 519
DB 284 TGATCTTGAGAGCTGTGCTCC 303

```

```

RESULT 31
US-09-918-995-38039
; Sequence 38039, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:

```

```
US-09-960-643-1
; Sequence 1, Application US/09960643
; Publication No. US2002008801A1
; GENERAL INFORMATION:
; APPLICANT: Yoganathan, Thillainathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: CAM-XI and its Uses
; FILE REFERENCE: KINE024
; CURRENT APPLICATION NUMBER: US/09/960,643
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2447
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(1498)
US-09-960-643-1

Query Match      1.9%; Score 21; DB 13; Length 2447;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      919 CATGGCGCTCATCACCCTACAT 939
DB      681 CATGGCGCTCATCACCCTACAT 701

RESULT 26
US-10-220-120-29
; Sequence 29, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Amlisa L.
; APPLICANT: LIT, Tommy F.
; APPLICANT: ROSEBERY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220,120
```

```
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
; 60/205,324; 60/205,286;
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;
; 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-15;
; 2000-05-16; 2000-05-17; 2000-05-17; 2000-05-17; 2000-05-17;
; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2689
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:333138.2:2000MAY01
US-10-220-120-29

Query Match      1.9%; Score 21; DB 13; Length 2689;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      919 CATGGCGCTCATCACCCTACAT 939
DB      848 CATGGCGCTCATCACCCTACAT 868

RESULT 27
US-09-935-464-1
; Sequence 1, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: PARKER, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 157875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-1

Query Match      1.9%; Score 21; DB 10; Length 157875;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      919 CATGGCGCTCATCACCCTACAT 939
DB      142369 CATGGCGCTCATCACCCTACAT 142389

RESULT 28
US-10-125-835-1
; Sequence 1, Application US/10125835
; Publication No. US20030092019A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
```


APPLICANT: Meyer, Joanne
 APPLICANT: Barrington-Martin, Rory
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING
 TITLE OF INVENTION: NEUROPSYCHIATRIC
 TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
 FILE REFERENCE: 3322/0H702 US0
 CURRENT APPLICATION NUMBER: US/10/125,835
 PRIOR FILING DATE: 2002-04-19
 PRIOR APPLICATION NUMBER: US/09/757,300
 PRIOR FILING DATE: 2001-01-09
 NUMBER OF SEQ ID NOS: 50
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 2
 LENGTH: 1383
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-125-835-2

Query Match 1.9%; Score 21; DB 15; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCCTACCTACAT 939
 Db 612 CATGGCGTCATCCTACCTACAT 632

RESULT 19
 US-10-152-319A-1918
 Sequence 1918, Application US/10152319A
 Publication No. US20040072160A1
 GENERAL INFORMATION:
 APPLICANT: Mendrick, Donna
 APPLICANT: Porter, Mark
 APPLICANT: Johnson, Kory
 APPLICANT: Higgs, Brandon
 APPLICANT: Castle, Arthur
 APPLICANT: Elashoff, Michael
 TITLE OF INVENTION: Molecular Toxicology Modeling
 FILE REFERENCE: 44921-5089-US
 CURRENT APPLICATION NUMBER: US/10/152,319A
 PRIOR FILING DATE: 2002-05-22
 PRIOR APPLICATION NUMBER: US 60/292,335
 PRIOR FILING DATE: 2001-05-22
 PRIOR APPLICATION NUMBER: US 60/297,523
 PRIOR FILING DATE: 2001-06-13
 PRIOR APPLICATION NUMBER: US 60/298,925
 PRIOR FILING DATE: 2001-06-19
 PRIOR APPLICATION NUMBER: US 60/303,810
 PRIOR FILING DATE: 2001-07-10
 PRIOR APPLICATION NUMBER: US 60/303,807
 PRIOR FILING DATE: 2001-07-10
 PRIOR APPLICATION NUMBER: US 60/303,808
 PRIOR FILING DATE: 2001-07-10
 PRIOR APPLICATION NUMBER: US 60/315,047
 PRIOR FILING DATE: 2001-08-28
 PRIOR APPLICATION NUMBER: US 60/324,928
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: US 60/330,867
 PRIOR FILING DATE: 2001-11-01
 PRIOR APPLICATION NUMBER: US 60/330,462
 PRIOR FILING DATE: 2001-10-22
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 2221
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1918
 LENGTH: 1388
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 OTHER INFORMATION: Genbank Accession No. NM_031573
 US-10-152-319A-1918

Query Match 1.9%; Score 21; DB 12; Length 1388;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGAGACATC 662
 Db 524 GACCTGAAGCCGAGACATC 544

RESULT 20
 US-09-935-464-4
 Sequence 4, Application US/09935464
 Publication No. US20030027153A1
 GENERAL INFORMATION:
 APPLICANT: Meyer, Joanne
 APPLICANT: Barrington-Martin, Rory
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIAT
 TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
 FILE REFERENCE: 3322/1H702 US1
 CURRENT APPLICATION NUMBER: US/09/935,464
 PRIOR FILING DATE: 2001-08-23
 PRIOR APPLICATION NUMBER: US 09/757,300
 PRIOR FILING DATE: 2001-01-09
 NUMBER OF SEQ ID NOS: 90
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 4
 LENGTH: 1738
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-935-464-4

Query Match 1.9%; Score 21; DB 10; Length 1738;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCCTACCTACAT 939
 Db 629 CATGGCGTCATCCTACCTACAT 649

RESULT 21
 US-10-125-835-4
 Sequence 4, Application US/10125835
 Publication No. US20030092019A1
 GENERAL INFORMATION:
 APPLICANT: Meyer, Joanne
 APPLICANT: Barrington-Martin, Rory
 APPLICANT: Parker, Alexander
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING
 TITLE OF INVENTION: NEUROPSYCHIATRIC
 FILE REFERENCE: 3322/0H702 US0
 CURRENT APPLICATION NUMBER: US/10/125,835
 PRIOR FILING DATE: 2002-04-19
 PRIOR APPLICATION NUMBER: US/09/757,300
 PRIOR FILING DATE: 2001-01-09
 NUMBER OF SEQ ID NOS: 50
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 4
 LENGTH: 1738
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-125-835-4

Query Match 1.9%; Score 21; DB 15; Length 1738;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCCTACCTACAT 939
 Db 629 CATGGCGTCATCCTACCTACAT 649

```
RESULT 22
US-10-037-270-809
; Sequence 809, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aigong J.
; APPLICANT: Yang, Yongshong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunwei
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillingshast, John
; APPLICANT: Dzmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP23
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 809
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1704)
; NAME/KEY: misc_feature
; LOCATION: (1)..(2165)
; OTHER INFORMATION: n = a,t,c or g
US-10-037-270-809

Query Match          1.9%; Score 21; DB 15; Length 2165;
Best Local Similarity 100.0%; Pred.No.1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          919 CATGGCGTCATCCTACAT 939
Db          885 CATGGCGTCATCCTACAT 905

RESULT 23
US-10-117-722-809
; Sequence 809, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dzmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
```

```
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 809
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1704)
; NAME/KEY: misc_feature
; LOCATION: (1)..(2165)
; OTHER INFORMATION: n = a,t,c or g
US-10-117-722-809

Query Match          1.9%; Score 21; DB 16; Length 2165;
Best Local Similarity 100.0%; Pred.No.1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          919 CATGGCGTCATCCTACAT 939
Db          885 CATGGCGTCATCCTACAT 905

RESULT 24
US-10-380-235-3
; Sequence 3, Application US/10380235
; Publication No. US20040072184A1
; GENERAL INFORMATION:
; APPLICANT: Yoganathan, Thillanathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: CANCER ASSOCIATED PROTEIN KINASES AND
; TITLE OF INVENTION: THEIR USES
; FILE REFERENCE: KINE-024CIP
; CURRENT APPLICATION NUMBER: US/10/380,235
; PRIOR FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: IB01/02237
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/250,555
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/233,999
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/237,419
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,423
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/238,558
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)...(1498)
US-10-380-235-3

Query Match          1.9%; Score 21; DB 12; Length 2447;
Best Local Similarity 100.0%; Pred.No.1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          919 CATGGCGTCATCCTACAT 939
Db          681 CATGGCGTCATCCTACAT 701

RESULT 25
```

Thu Jul 1 09:50:16 2004

us-10-083-641a-6.rnpb

Page 12

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 976
; LENGTH: 2055
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097948_FLI
US-10-425-114-976

Query Match          1.8%; Score 20; DB 13; Length 2055;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 518 CCGGGCGGCACTTTTCGAC 537
Db 416 CCGGGCGGCACTTTTCGAC 435

RESULT 36
US-10-425-114-5579
; Sequence 5579, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowallik, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 5579
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700465007_FLI
US-10-425-114-5579

Query Match          1.8%; Score 20; DB 13; Length 2382;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 518 CCGGGCGGCACTTTTCGAC 537
Db 686 CCGGGCGGCACTTTTCGAC 705

RESULT 37
US-09-918-995-10392
; Sequence 10392, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseed, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918, 995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235, 076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 10392
```

```

; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(456)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10392

Query Match          1.7%; Score 19; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 648 AACCCGAGACATCATGT 666
Db 232 AACCCGAGACATCATGT 250

RESULT 38
US-10-027-632-89093/C
; Sequence 89093, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 89093
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-89093

Query Match          1.7%; Score 19; DB 13; Length 504;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 CAGAGGCTGGGTACAG 241
Db 275 CAGAGGCTGGGTACAG 257

RESULT 39
US-10-027-632-89093/C
; Sequence 89093, Application US/10027632
; Publication No. US20030204075A3
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```

```

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 38039
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-38039

Query Match
Best Local Similarity 1.8%; Score 20; DB 10; Length 493;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 923 GGGCTATCATCCTACATCCT 942
DB 471 GGGCTATCATCCTACATCCT 490

RESULT 32
US-10-029-386-9853
; Sequence 9853, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Handel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 9853
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR15.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA858002.1, EVALUE 1.00e-111
; OTHER INFORMATION: NT HIT: G16162072, EVALUE 1.00e-123
; OTHER INFORMATION: SWISSPROT HIT: P53355, EVALUE 5.00e-30
US-10-029-386-9853

Query Match
Best Local Similarity 1.8%; Score 20; DB 15; Length 535;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 AGCCGGCGGGGTGTGAGCCG 401
DB 248 AGCCGGCGGGGTGTGAGCCG 267

RESULT 33
US-10-087-192-1835
; Sequence 1835, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1835
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1835

Query Match
Best Local Similarity 1.8%; Score 20; DB 13; Length 1716;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 AGCCGGCGGGGTGTGAGCCG 401
DB 190 AGCCGGCGGGGTGTGAGCCG 209

RESULT 34
US-10-425-114-25191
; Sequence 25191, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 25191
; LENGTH: 1934
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3732-006-E7_FLI
US-10-425-114-25191

Query Match
Best Local Similarity 1.8%; Score 20; DB 13; Length 1934;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 518 CCGGCGGGAACCTTTGAC 537
DB 526 CCGGCGGGAACCTTTGAC 545

RESULT 35
US-10-425-114-976
; Sequence 976, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
```

PRIOR FILING DATE: 2002-05-06
 NUMBER OF SEQ ID NOS: 1140
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 230
 LENGTH: 2287
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 FEATURE:
 OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_031985
 US-10-191-803-230

Query Match 2.2%; Score 24; DB 16; Length 2287;
 Best Local Similarity 100.0%; Pred.No. 0.045;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGAGACCTCATG 665
 DB 673 GACCTGAAGCCGAGACCTCATG 696

RESULT 11
 US-09-918-995-313
 Sequence 313, Application US/09918995
 Publication No. US20030073623A1
 GENERAL INFORMATION:
 APPLICANT: Hyseq, Inc.
 TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 FILE REFERENCE: 20411-756
 CURRENT APPLICATION NUMBER: US/09/918,995
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: US/09/235,076
 PRIOR FILING DATE: 1999-01-20
 NUMBER OF SEQ ID NOS: 38054
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 313
 LENGTH: 488
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (1)...(488)
 OTHER INFORMATION: n = A,T,C or G
 US-09-918-995-313

Query Match 1.9%; Score 21; DB 10; Length 488;
 Best Local Similarity 100.0%; Pred.No. 2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCACCCTACAT 939
 DB 185 CATGGCGTCATCACCCTACAT 205

RESULT 12
 US-10-027-632-136289
 Sequence 136289, Application US/10027632
 Publication No. US20020198371A1
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827.129
 CURRENT APPLICATION NUMBER: US/10/027,632
 PRIOR FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/218,006
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: US 60/156,358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146,002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 136289
 LENGTH: 692
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-136289

Query Match 1.9%; Score 21; DB 13; Length 692;
 Best Local Similarity 100.0%; Pred.No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCACCCTACAT 939
 DB 645 CATGGCGTCATCACCCTACAT 665

RESULT 13
 US-10-027-632-136290
 Sequence 136290, Application US/10027632
 Publication No. US20020198371A1
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827.129
 CURRENT APPLICATION NUMBER: US/10/027,632
 PRIOR FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/218,006
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/167,363
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: US 60/156,358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146,002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 136290
 LENGTH: 692
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-136290

Query Match 1.9%; Score 21; DB 13; Length 692;
 Best Local Similarity 100.0%; Pred.No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCACCCTACAT 939
 DB 645 CATGGCGTCATCACCCTACAT 665

RESULT 14
 US-10-027-632-136289
 Sequence 136289, Application US/10027632
 Publication No. US20030204075A9
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827.129
 CURRENT APPLICATION NUMBER: US/10/027,632
 PRIOR FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/218,006
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24

FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 136289
LENGTH: 692
TYPE: DNA
ORGANISM: Human
US-10-027-632-136289

Query Match 1.9%; Score 21; DB 16; Length 692;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCACCCTACAT 939
DB 645 CATGGCGTCATCACCCTACAT 665

RESULT 15
US-10-027-632-136290
Sequence 136290, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 136290
LENGTH: 692
TYPE: DNA
ORGANISM: Human
US-10-027-632-136290

Query Match 1.9%; Score 21; DB 16; Length 692;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 919 CATGGCGTCATCACCCTACAT 939
DB 645 CATGGCGTCATCACCCTACAT 665

DB 645 CATGGCGTCATCACCCTACAT 665

RESULT 16
US-10-369-493-42833
Sequence 42833, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianteng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 42833
LENGTH: 1052
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-10-369-493-42833

Query Match 1.9%; Score 21; DB 16; Length 1052;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGGAGAACATC 662
DB 391 GACCTGAAGCCGGAGAACATC 411

RESULT 17
US-09-935-464-2
Sequence 2, Application US/09935464
Publication No. US20030027153A1
GENERAL INFORMATION:
APPLICANT: Meyer, Joanne
APPLICANT: Barrington-Martin, Rory
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
FILE REFERENCE: 3322/1H702 US1
CURRENT APPLICATION NUMBER: US/09/935,464
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: US 09/757,300
PRIOR FILING DATE: 2001-01-09
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapiens
US-09-935-464-2

Query Match 1.9%; Score 21; DB 10; Length 1383;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCACCCTACAT 939
DB 612 CATGGCGTCATCACCCTACAT 632

RESULT 18
US-10-125-835-2
Sequence 2, Application US/10125835
Publication No. US20030092019A1
GENERAL INFORMATION:

```
RESULT 42
US-10-282-122A-33150
; Sequence 33150, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33150
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Pseudomonas syringae
US-10-282-122A-33150

Query Match 1.7%; Score 19; DB 13; Length 1227;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 GACAGCATGCAGCCAGCC 690
Db 817 GACAGCATGCAGCCAGCC 835

RESULT 43
US-10-425-114-16427
; Sequence 16427, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(52313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16427
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3064-097-F3_FLI
US-10-425-114-16427

Query Match 1.7%; Score 19; DB 13; Length 1467;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 648 AAGCCGGAGAACATCATGT 666
Db 421 AAGCCGGAGAACATCATGT 439

RESULT 44
US-10-084-817-78
; Sequence 78, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 78
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 471362.33
US-10-084-817-78

Query Match 1.7%; Score 19; DB 15; Length 1640;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 648 AAGCCGGAGAACATCATGT 666
Db 234 AAGCCGGAGAACATCATGT 252

RESULT 45
US-10-210-120-25
; Sequence 25, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UN-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
```

PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/167,363
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: US 60/156,358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146,002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 89093
 LENGTH: 504
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-89093

Query Match 1.7%; Score 19; DB 16; Length 504;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 CAGAGGCTGGGTACAGAG 241
 Db 275 CAGAGGCTGGGTACAGAG 257

RESULT 40
 US-09-864-761-8409/c
 Sequence 8409, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aecomics-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 8409
 LENGTH: 559
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC004551.1
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
 OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 3.3
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
 US-09-864-761-8409

Query Match 1.7%; Score 19; DB 9; Length 559;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 CTGCCTTCCTCTGGGTCTG 262
 Db 306 CTGCCTTCCTCTGGGTCTG 288

RESULT 41
 US-10-260-238-160
 Sequence 160, Application US/10260238
 Publication No. US20040016025A1
 GENERAL INFORMATION:
 APPLICANT: Budworth, Paul R.
 APPLICANT: Moughamer, Todd G.
 APPLICANT: Briggs, Steven P.
 APPLICANT: Cooper, Bret
 APPLICANT: Glazebrook, Jane
 APPLICANT: Goff, Stephen A.
 APPLICANT: Katagiri, Fumiyaki
 APPLICANT: Kreps, Joel
 APPLICANT: Provart, Nicholas
 APPLICANT: Ricke, Darrell
 APPLICANT: Zhu, Tong
 TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
 FILE REFERENCE: 60111-NP
 CURRENT APPLICATION NUMBER: US/10/260,238
 CURRENT FILING DATE: 2002-09-26
 PRIOR APPLICATION NUMBER: US 60/325,448
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: US 60/325,277
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: US 60/370,620
 PRIOR FILING DATE: 2002-04-04
 NUMBER OF SEQ ID NOS: 6077
 SEQ ID NO 160
 LENGTH: 1145
 TYPE: DNA
 ORGANISM: Oryza sativa
 US-10-260-238-160

Query Match 1.7%; Score 19; DB 16; Length 1145;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 641 TGACCTGAAGCCGGAGAAC 659
 Db 341 TGACCTGAAGCCGGAGAAC 359

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2004, 18:50:10 ; Search time 538 Seconds

(without alignments)
9340.072 Million cell updates/sec

Title: US-10-083-641A-6

Perfect score: 1093
Sequence: 1 gntatgmatcggttatac.....ccaccangtsgaccatnc 1093

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3031105 seqs, 2298700234 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6062210

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCRT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCRTS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1093	100.0	1093	US-10-083-641A-6	Sequence 6, Appl1
2	44	4.0	65	US-09-908-975-1053	Sequence 1053, Ap
3	26	2.4	1592	US-10-087-192-1832	Sequence 1832, Ap
4	26	2.4	2132	US-10-262-445-118	Sequence 118, Ap
5	26	2.4	2224	US-10-055-585-111	Sequence 11, Appl
6	26	2.4	126413	US-10-087-192-1831	Sequence 1831, Ap
7	26	2.4	152330	US-10-087-192-1834	Sequence 1834, Ap
8	24	2.2	1800	US-10-316-253-31	Sequence 31, Appl
9	24	2.2	2287	US-10-316-253-33	Sequence 33, Appl
10	24	2.2	2287	US-10-191-803-330	Sequence 230, App
11	21	1.9	488	US-09-918-995-313	Sequence 313, App
12	21	1.9	692	US-10-027-632-136289	Sequence 136289,
13	21	1.9	692	US-10-027-632-136290	Sequence 136290,
14	21	1.9	692	US-10-027-632-136289	Sequence 136289,

15	21	1.9	692	US-10-027-632-136290	Sequence 136290,
16	21	1.9	1052	US-10-369-493-42833	Sequence 42833, A
17	21	1.9	1383	US-09-935-464-2	Sequence 2, Appl1
18	21	1.9	1383	US-10-125-835-2	Sequence 4, Appl1
19	21	1.9	1388	US-10-152-319A-1918	Sequence 1918, Ap
20	21	1.9	1738	US-09-935-464-4	Sequence 4, Appl1
21	21	1.9	1738	US-10-125-835-4	Sequence 4, Appl1
22	21	1.9	2165	US-10-037-270-809	Sequence 809, App
23	21	1.9	2165	US-10-117-722-809	Sequence 809, App
24	21	1.9	2447	US-10-380-235-3	Sequence 1, Appl1
25	21	1.9	2447	US-09-960-643-1	Sequence 1, Appl1
26	21	1.9	2669	US-10-220-120-29	Sequence 29, Appl
27	21	1.9	157875	US-09-935-464-1	Sequence 1, Appl1
28	21	1.9	157875	US-10-125-835-1	Sequence 1, Appl1
29	20	1.8	147	US-10-029-385-23553	Sequence 23553, A
30	20	1.8	447	US-09-664-761-11330	Sequence 11330, A
31	20	1.8	493	US-09-918-995-38039	Sequence 38039, A
32	20	1.8	535	US-10-029-386-9853	Sequence 9853, Ap
33	20	1.8	1716	US-10-087-192-1835	Sequence 1835, Ap
34	20	1.8	1934	US-10-425-114-25191	Sequence 25191, A
35	20	1.8	2055	US-10-425-114-976	Sequence 976, App
36	20	1.8	2382	US-10-425-114-5579	Sequence 5579, Ap
37	19	1.7	456	US-09-918-995-10392	Sequence 10392, A
38	19	1.7	504	US-10-027-632-89093	Sequence 89093, A
39	19	1.7	504	US-10-027-632-89093	Sequence 89093, A
40	19	1.7	559	US-09-664-761-8409	Sequence 8409, Ap
41	19	1.7	1145	US-10-260-238-160	Sequence 160, App
42	19	1.7	1227	US-10-282-122A-33150	Sequence 33150, A
43	19	1.7	1467	US-10-425-114-16427	Sequence 16427, A
44	19	1.7	1640	US-10-084-817-78	Sequence 78, Appl
45	19	1.7	3164	US-10-210-120-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-10-083-641A-6
Sequence 6, Application US/10083641A
Publication No. US20030017568A1
GENERAL INFORMATION: SMOOTH MUSCLE MYOSIN PHOSPHATASE ASSOCIATED KINASE
APPLICANT: HAYSTED, TIMOTHY A
TITLE OF INVENTION: SMOOTH MUSCLE MYOSIN PHOSPHATASE ASSOCIATED KINASE
FILE REFERENCE: 1579-647
CURRENT APPLICATION NUMBER: US/10/083,641A
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/271,436
PRIOR FILING DATE: 2001-02-27
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 1093
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: Putative nucleotide
OTHER INFORMATION: sequence of smooth muscle MYPT-Kinase
FEATURE:
NAME/KEY: Unsure
LOCATION: (2), (7), (39), (1056), (1081) and (1092)
OTHER INFORMATION: N can be A, C, G or T
US-10-083-641A-6
Query Match 100.0%; Score 1093; DB 13; Length 1093;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1093; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GNTATGATATCGGTTAATCGCGAGCTCCGCGGAGCTGAGTCCCTCTCA 60
Db 1 GNTATGATATCGGTTAATCGCGAGCTCCGCGGAGCTGAGTCCCTCTCA 60
OY 61 GACCTCCTTTCTTCGCGCTCAGACGAGATTAACTTCACTTGACTGTCTTGAGTCC 120

Db 61 GACCTCTCTTCTTCTGCCCCGACGAGGAGATTAACTCACTTGACTGTCTTGGGCTCC 120
Qy 121 CGGTGCGGGGCGAGCGGTCTCTCTCTCAAGGCAATCCCACTGTCTCATGAGCTCT 180
Db 121 CGGTGCGGGGCGAGCGGTCTCTCTCTCAAGGCAATCCCACTGTCTCATGAGCTCT 180
Qy 181 TTGGGGAGTCTGTGTGTGTGGGAACTGGGAAAGATGACAGAGGCTGGGGTACAGA 240
Db 181 TTGGGGAGTCTGTGTGTGTGGGAACTGGGAAAGATGACAGAGGCTGGGGTACAGA 240
Qy 241 GTCTGCTCTCTCTCTGTGGGTCTGAGCGCTTAAGTCTTCTTCCCAACAGCGGCTTC 300
Db 241 GTCTGCTCTCTCTCTGTGGGTCTGAGCGCTTAAGTCTTCTTCCCAACAGCGGCTTC 300
Qy 301 GGCATGTGTGGGAGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 GGCATGTGTGGGAGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy 361 AAGAGCGGCGCTGCGCTGCTGAGCGCGGCTGAGCGGCTGAGGAGGAGGAGGAGG 420
Db 361 AAGAGCGGCGCTGCGCTGCTGAGCGCGGCTGAGCGGCTGAGGAGGAGGAGGAGG 420
Qy 421 GTGAGCATCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 421 GTGAGCATCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Qy 481 AACAGAGCAGATGT 540
Db 481 AACAGAGCAGATGT 540
Qy 541 CTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 541 CTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Qy 601 ACCGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 ACCGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 TCATGT 720
Db 661 TCATGT 720
Qy 721 CGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 721 CGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 781 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Qy 841 TGCTCTCAACATCT 900
Db 841 TGCTCTCAACATCT 900
Qy 901 GGAAGCTGATATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 901 GGAAGCTGATATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Qy 961 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 961 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Qy 1021 TGCAAACTGTGGGCTGAGAGATGTCTCTGGGAACTGTGATGTCTGGGCCCCCA 1080
Db 1021 TGCAAACTGTGGGCTGAGAGATGTCTCTGGGAACTGTGATGTCTGGGCCCCCA 1080
Qy 1081 NGTAGGACCATNC 1093
Db 1081 NGTAGGACCATNC 1093

RESULT 2
US-09-908-975-3053

Sequence 3053, Application US/0908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: MASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: MINTZ, Eli
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
FILE REFERENCE: 3698-0005
CURRENT FILING DATE: US/09/08, 975
PRIOR FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: Patent version 3.0
SEQ ID NO 3053
LENGTH: 65
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-908-975-3053

Query Match
Best Local Similarity 100.0%; Score 44; DB 10; Length 65;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 723 CACGAGTCGAGGCGCGGTAGCGAGTTCAGAGATCTTTGGCAG 766
Db 1 CACGAGTCGAGGCGCGGTAGCGAGTTCAGAGATCTTTGGCAG 44

RESULT 3
US-10-087-192-1832
Sequence 1832, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 52945200122
CURRENT FILING DATE: 2002-03-01
CURRENT FILING DATE: 2002-03-01
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 09/798,586
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1832
LENGTH: 1592
TYPE: DNA
ORGANISM: Mus musculus
US-10-087-192-1832

Query Match
Best Local Similarity 100.0%; Score 26; DB 13; Length 1592;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 445 CACCCCAACATCATCAGCTGACAGA 470
Db 216 CACCCCAACATCATCAGCTGACAGA 241

RESULT 4
US-10-262-445-118
Sequence 118, Application US/10262445
Publication No. US20040014058A1
GENERAL INFORMATION:
APPLICANT: Alsbjork II, John

Thu Jul 1 09:50:16 2004

us-10-083-641a-6.rnpb

LOCATION: (1)...(126413)
OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1831

Query Match 2.4% Score 26; DB 13; Length 126413;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 445 CACCCCAACATCATCAGCTGCACGA 470
DB 65016 CACCCCAACATCATCAGCTGCACGA 65041

RESULT 7
US-10-087-192-1834

Sequence 1834, Application US/10087192
Publication No. US20020182586A1

GENERAL INFORMATION:

APPLICANT: Morris, David W.

APPLICANT: Engelhard, Eric K.

TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

FILE REFERENCE: 529452000122

CURRENT APPLICATION NUMBER: US/10/087,192

PRIOR FILING DATE: 2002-03-01

PRIOR APPLICATION NUMBER: US 09/747,377

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: US 09/798,586

NUMBER OF SEQ ID NOS: 2059

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1834

LENGTH: 152330

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(152330)

OTHER INFORMATION: n = A,T,C or G

US-10-087-192-1834

Query Match 2.4% Score 26; DB 13; Length 152330;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 923 GACGTCATCAGCTCATCTGTGAGT 948
DB 120680 GACGTCATCAGCTCATCTGTGAGT 120705

RESULT 8
US-10-316-253-31

Sequence 31, Application US/10316253
Publication No. US20030162706A1

GENERAL INFORMATION:

APPLICANT: The Procter & Gamble Company

APPLICANT: Peters, Kevin

APPLICANT: Thompson, Larry

APPLICANT: Wang, Feng

APPLICANT: Greis, Kenneth

TITLE OF INVENTION: Angiogenesis Modulating Proteins

FILE REFERENCE: 8665M

CURRENT APPLICATION NUMBER: US/10/316,253

PRIOR FILING DATE: 2002-12-10

PRIOR APPLICATION NUMBER: US 60/355,295

NUMBER OF SEQ ID NOS: 308

SOFTWARE: PatentIn version 3.1

SEQ ID NO 31

LENGTH: 1800

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

US-10-316-253-31

NAME/KEY: CDS
LOCATION: (134)...(1642)
OTHER INFORMATION:
US-10-316-253-31

Query Match 2.2% Score 24; DB 15; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGAGAACATCATG 665
DB 716 GACCTGAAGCCGAGAACATCATG 739

RESULT 9
US-10-316-253-33

Sequence 33, Application US/10316253
Publication No. US20030162706A1

GENERAL INFORMATION:

APPLICANT: The Procter & Gamble Company

APPLICANT: Peters, Kevin

APPLICANT: Thompson, Larry

APPLICANT: Wang, Feng

APPLICANT: Greis, Kenneth

TITLE OF INVENTION: Angiogenesis Modulating Proteins

FILE REFERENCE: 8665M

CURRENT APPLICATION NUMBER: US/10/316,253

PRIOR FILING DATE: 2002-12-10

PRIOR APPLICATION NUMBER: US 60/355,295

NUMBER OF SEQ ID NOS: 308

SOFTWARE: PatentIn version 3.1

SEQ ID NO 33

LENGTH: 2287

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

NAME/KEY: CDS

LOCATION: (22)...(1599)

OTHER INFORMATION:

US-10-316-253-33

Query Match 2.2% Score 24; DB 15; Length 2287;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGAGAACATCATG 665
DB 673 GACCTGAAGCCGAGAACATCATG 696

RESULT 10
US-10-191-803-230

Sequence 230, Application US/10191803
Publication No. US20040014040A1

GENERAL INFORMATION:

APPLICANT: MENDRICK, Donna

APPLICANT: PORTER, Mark

APPLICANT: JOHNSON, Kory

APPLICANT: HIGGS, Brandon

APPLICANT: CASTLE, Arthur

APPLICANT: ELASHOFF, Michael

TITLE OF INVENTION: Cardiolipin Molecular Toxicology Modeling

FILE REFERENCE: 44921-5090US

CURRENT APPLICATION NUMBER: US/10/191,803

PRIOR FILING DATE: 2002-07-10

PRIOR APPLICATION NUMBER: US 60/303,819

PRIOR FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: US 60/305,623

PRIOR FILING DATE: 2001-07-17

PRIOR APPLICATION NUMBER: US 60/369,351

PRIOR FILING DATE: 2002-04-03

PRIOR APPLICATION NUMBER: US 60/377,611

US-10-191-803-230

```

; APPLICANT: Burgess, Catherine
; APPLICANT: Carterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Giot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patlurajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Tausper Jr., Raymond J.
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-462D
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Cytaseqlist version 0.1
; SEQ ID NO 118
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)..(1455)
US-10-262-445-118

Query Match 2.4%; Score 26; DB 16; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CAGTTCCTCAAGCAGATCCTGACCG 604
Db 445 CAGTTCCTCAAGCAGATCCTGACCG 470

RESULT 5
US-10-059-585-11
; Sequence 11, Application US/10059585
; Publication No. US2003008276A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
```

```

; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ki
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; FILE REFERENCE: 06501-096001
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(1576)
US-10-059-585-11

Query Match 2.4%; Score 26; DB 15; Length 2224;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CAGTTCCTCAAGCAGATCCTGACCG 604
Db 566 CAGTTCCTCAAGCAGATCCTGACCG 591

RESULT 6
US-10-087-192-1831
; Sequence 1831, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1831
; LENGTH: 126413
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
```

Thu Jul 1 09:50:16 2004

us-10-083-641a-6.rnpb

Page 15

; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 25
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-25

Query Match 1.7%; Score 19; DB 15; Length 3164;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 648 AAGCCGAGAACATCATGT 666
Db 2118 AAGCCGAGAACATCATGT 2136

Search completed: June 30, 2004, 21:25:05
Job time : 540 secs

Thu Jul 1 09:50:15 2004

us-10-083-641a-6.rnt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2004, 18:54:05 ; Search time 99 Seconds
(without alignments)
6126.882 Million cell updates/sec

Title: US-10-083-641A-6
Perfect score: 1093
Sequence: 1 gntatgatacgcgttcaat.....ccaccangtaggaccatnc 1093

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/6C_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	4.0	1429	2	US-09-159-385-4
2	44	4.0	1429	3	US-09-186-277-4
3	26	2.4	2132	2	US-09-159-385-3
4	26	2.4	2132	3	US-09-186-277-3
5	21	1.9	2165	4	US-09-620-312D-809
6	20	1.8	295	1	US-09-313-294A-6448
7	20	1.8	748	1	US-08-510-878-3
8	20	1.8	28720	4	US-09-341-587-7
9	18	1.6	558	4	US-09-252-991A-4654
10	18	1.6	1335	4	US-09-252-991A-4449
11	18	1.6	1347	4	US-09-252-991A-4158
12	18	1.6	1619	4	US-09-394-455-14
13	18	1.6	3051	4	US-09-409-604-1
14	17	1.6	20	4	US-09-980-052-155
15	17	1.6	261	4	US-09-107-532A-1941
16	17	1.6	375	4	US-09-621-976-8505
17	17	1.6	411	4	US-09-252-991A-8200
18	17	1.6	467	4	US-09-621-976-2248
19	17	1.6	507	4	US-09-252-991A-13127
20	17	1.6	706	3	US-08-797-812-25
21	17	1.6	843	4	US-09-252-991A-11363
22	17	1.6	930	4	US-09-252-991A-11343
23	17	1.6	984	4	US-09-252-991A-2244
24	17	1.6	993	4	US-09-252-991A-11960
25	17	1.6	999	4	US-09-252-991A-2600
26	17	1.6	999	4	US-09-252-991A-8255
27	17	1.6	1008	4	US-09-394-455-3

28	17	1.6	1014	4	US-09-252-991A-12024	Sequence 12024, A
29	17	1.6	1095	4	US-09-252-991A-2312	Sequence 2312, Ap
30	17	1.6	1119	4	US-09-252-991A-11990	Sequence 11990, A
31	17	1.6	1134	4	US-09-252-991A-8104	Sequence 8104, Ap
32	17	1.6	1281	4	US-09-107-532A-3059	Sequence 3059, Ap
33	17	1.6	1290	4	US-09-252-991A-12508	Sequence 12508, A
34	17	1.6	1307	4	US-09-023-655-293	Sequence 293, App
35	17	1.6	1314	4	US-09-252-991A-8001	Sequence 8001, Ap
36	17	1.6	1431	4	US-09-252-991A-2821	Sequence 2821, Ap
37	17	1.6	1450	4	US-09-620-312D-1058	Sequence 1058, Ap
38	17	1.6	1755	4	US-09-252-991A-8133	Sequence 8133, Ap
39	17	1.6	1815	3	US-09-178-252-3	Sequence 3, Appli
40	17	1.6	1815	3	US-09-178-252-7	Sequence 7, Appli
41	17	1.6	1815	3	US-09-178-252-8	Sequence 8, Appli
42	17	1.6	1815	4	US-09-826-660-3	Sequence 3, Appli
43	17	1.6	1815	4	US-09-826-660-7	Sequence 7, Appli
44	17	1.6	1815	4	US-09-826-660-8	Sequence 8, Appli
45	17	1.6	1862	4	US-09-336-643A-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-09-159-385-4
Sequence 4, Application US/09159385
Patent No. 5958748
GENERAL INFORMATION:
APPLICANT: AKIRA, SHIZUO
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
FILE REFERENCE: PH-569
CURRENT APPLICATION NUMBER: US/09/159,385
EARLIER APPLICATION NUMBER: JP97/261589
EARLIER FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1429
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (10) ..(1353)
US-09-159-385-4

Query Match 4.0%; Score 44; DB 2; Length 1429;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CCGAGGTGAGCATCTGCGCGAGATCCGCCACCCCAACATCAT 458
DB 196 CCGAGGTGAGCATCTGCGCGAGATCCGCCACCCCAACATCAT 239

RESULT 2
US-09-186-277-4
Sequence 4, Application US/09186277
Patent No. 6171841
GENERAL INFORMATION:
APPLICANT: AKIRA, SHIZUO
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
FILE REFERENCE: 081356/0128
CURRENT APPLICATION NUMBER: US/09/186,277
EARLIER FILING DATE: 1998-11-05
EARLIER APPLICATION NUMBER: JP97/261589
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1429

TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (10)..(1353)
US-09-186-277-4

Query Match
Best Local Similarity 4.0%; Score 44; DB 3; Length 1429;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGTGAGCATCTCGCGGAGATCCGCCACCCCAACATCAT 458
DB 196 CGCGAGTGAGCATCTCGCGGAGATCCGCCACCCCAACATCAT 239

RESULT 3
US-09-159-385-3
Sequence 3, Application US/09159385
Patent No. 5958748
GENERAL INFORMATION:
APPLICANT: AKIRA, SHIZUO
APPLICANT: KAWAI, TARO
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
FILE REFERENCE: PH-569
CURRENT APPLICATION NUMBER: US/09/159,385
CURRENT FILING DATE: 1998-09-23
EARLIER APPLICATION NUMBER: JP97/261589
EARLIER FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2132
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)..(1455)
US-09-159-385-3

Query Match
Best Local Similarity 2.4%; Score 26; DB 2; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGCAGATCCTGGACGG 604
DB 445 CAGTTCTCAAGCAGATCCTGGACGG 470

RESULT 4
US-09-186-277-3
Sequence 3, Application US/09186277
Patent No. 6171841
GENERAL INFORMATION:
APPLICANT: AKIRA, SHIZUO
APPLICANT: KAWAI, TARO
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
FILE REFERENCE: 081356/0128
CURRENT APPLICATION NUMBER: US/09/186,277
CURRENT FILING DATE: 1998-11-05
EARLIER APPLICATION NUMBER: JP97/261589
EARLIER FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2132
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)..(1455)
US-09-186-277-3

Query Match
Best Local Similarity 2.4%; Score 26; DB 3; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCTCAAGCAGATCCTGGACGG 604
DB 445 CAGTTCTCAAGCAGATCCTGGACGG 470

RESULT 5
US-09-620-312D-809
Sequence 809, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc FL_genes Version 1.0
SEQ ID NO 809
LENGTH: 2165
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1704)
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(2165)
OTHER INFORMATION: n = a,c,t,c or g
US-09-620-312D-809

Query Match
Best Local Similarity 1.9%; Score 21; DB 4; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATGGCGTCATCACCTTACAT 939
DB 885 CATGGCGTCATCACCTTACAT 905

RESULT 6
US-09-313-294A-6448
Sequence 6448, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Ialigudi, Ragunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 6448
LENGTH: 295
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6476212 700351821H1
LOCATION: 290
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6448

Query Match 1.8%; Score 20; DB 4; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 TCGCCACCCCAACATCATC 459
DB 157 TCGCCACCCCAACATCATC 176

RESULT 7
US-08-510-878-3
Sequence 3, Application US/08510878
Patent No. 576771
GENERAL INFORMATION:
APPLICANT: Yu, Fujio
TITLE OF INVENTION: A KANAMYCIN RESISTANCE GENE DERIVED FROM
TITLE OF INVENTION: MICROORGANISMS OF THE GENUS RHODOCOCUS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,878
FILING DATE: 03-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 1254-121
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 748 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-510-878-3

Query Match 1.8%; Score 20; DB 1; Length 748;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 ACATCATGTTGCTGACACAG 677
DB 708 ACATCATGTTGCTGACACAG 727

RESULT 8
US-09-341-587-7
Sequence 7, Application US/09341587
Patent No. 6346606
GENERAL INFORMATION:
APPLICANT: Mollenhauer, Jan
TITLE OF INVENTION: Protein Containing an SRCR Domain
FILE REFERENCE: 4121-108
CURRENT APPLICATION NUMBER: US/09/341,587
CURRENT FILING DATE: 1999-08-31
EARLIER APPLICATION NUMBER: PCT/DE98/00096
EARLIER FILING DATE: 1998-01-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 7
LENGTH: 28720
TYPE: DNA
ORGANISM: Homo sapiens
US-09-341-587-7

Query Match 1.8%; Score 20; DB 4; Length 28720;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 952 TGAGATGGGACAGGGGCTCA 971
DB 22565 TGAGATGGGACAGGGGCTCA 22584

RESULT 9
US-09-252-991A-4654/c
Sequence 4654, Application US/09252991A
Patent No. 6551785
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4654
LENGTH: 558
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4654

Query Match 1.6%; Score 18; DB 4; Length 558;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GGCACCGGCATGAGTAC 345
DB 381 GGCACCGGCATGAGTAC 364

RESULT 10
US-09-252-991A-4449/c
Sequence 4449, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

```
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 4449
/ LENGTH: 1335
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
/ US-09-252-991A-4449

Query Match
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GGCACCGCATGAGTAC 345
Db 491 GGCACCGCATGAGTAC 474

RESULT 11
US-09-252-991A-4158
/ Sequence 4158, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 4158
/ LENGTH: 1347
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
/ US-09-252-991A-4158

Query Match
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GGCACCGCATGAGTAC 345
Db 893 GGCACCGCATGAGTAC 910

RESULT 12
US-09-394-455-14
/ Sequence 14, Application US/09394455
/ Patent No. 6531305
/ GENERAL INFORMATION:
/ APPLICANT: Witman, George F.
/ APPLICANT: San Agustín, Jovenal
/ APPLICANT: Leszyk, John D.
/ TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
/ FILE REFERENCE: 07917/078001
/ CURRENT APPLICATION NUMBER: US/09/394,455
/ PRIOR FILING DATE: 1999-09-10
/ PRIOR APPLICATION NUMBER: US 60/099,771
/ PRIOR FILING DATE: 1998-09-10
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14
```

```
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Ovine
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (30)...(1058)
/ US-09-394-455-14

Query Match
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 GACCTGAAGCCGAGAAC 659
Db 504 GACCTGAAGCCGAGAAC 521

RESULT 13
US-09-409-604-1
/ Sequence 1, Application US/09409604A
/ Patent No. 6632934
/ GENERAL INFORMATION:
/ APPLICANT: Moreadith, Randall W.
/ APPLICANT: Zinn, Andrew R.
/ APPLICANT: Watson, Mark L.
/ APPLICANT: Inoue, No. 6632934mitsu
/ APPLICANT: Hesse, Karl D.
/ TITLE OF INVENTION: MORC GENE COMPOSITIONS AND METHODS OF USE
/ FILE REFERENCE: URSID:573
/ CURRENT APPLICATION NUMBER: US/09/409,604A
/ PRIOR FILING DATE: 1999-09-30
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 3051
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (105)..(2954)
/ US-09-409-604-1

Query Match
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 AGAGTCTGCTCTCTCT 255
Db 1578 AGAGTCTGCTCTCTCT 1595

RESULT 14
US-09-980-052-155
/ Sequence 155, Application US/09980052
/ Patent No. 6670130
/ GENERAL INFORMATION:
/ APPLICANT: KIM, Jeong Joon; SJ HIGHTECH Co., Ltd.
/ APPLICANT: Kim, Cheol Min
/ APPLICANT: Park, Hee Kyung
/ TITLE OF INVENTION: Oligonucleotide for detection and identification of Mycobacteria
/ FILE REFERENCE: PP05020/PCT
/ CURRENT APPLICATION NUMBER: US/09/980,052
/ PRIOR FILING DATE: 2001-11-28
/ PRIOR APPLICATION NUMBER: KR 10-1999-0019631
/ PRIOR FILING DATE: 1999-05-29
/ PRIOR APPLICATION NUMBER: KR 10-1999-0019632
/ PRIOR FILING DATE: 1999-05-29
/ PRIOR APPLICATION NUMBER: KR 10-1999-0019633
/ PRIOR FILING DATE: 1999-05-29
/ PRIOR APPLICATION NUMBER: KR 10-1999-0019634
/ PRIOR FILING DATE: 1999-05-29
/ PRIOR APPLICATION NUMBER: KR 10-1999-0019635
```

PRIOR FILING DATE: 1999-05-29
PRIOR APPLICATION NUMBER: KR 10-2000-0018189
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 243
SOFTWARE: Koparentin 1.71
SEQ ID NO 155
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: sequence of probe or primer for detecting Mycobacterium smegmatis
US-09-380-052-155

Query Match 1.6%; Score 17; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GTGTGAGCCGTGAGGAG 408
DB 1 GTGTGAGCCGTGAGGAG 17

RESULT 15
US-09-107-532A-1341
Sequence 1941, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-8277
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1941:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYBOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...261
SEQUENCE DESCRIPTION: SEQ ID NO: 1941:

US-09-107-532A-1941
Query Match 1.6%; Score 17; DB 4; Length 261;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 653 GGAGAACATCATGTTGC 669
DB 115 GGAGAACATCATGTTGC 131

RESULT 16
US-09-621-976-8505
Sequence 8505, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8505
LENGTH: 375
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 333,344,348
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-8505

Query Match 1.6%; Score 17; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GGGCAGCTGGACTCCCT 56
DB 118 GGGCAGCTGGACTCCCT 134

RESULT 17
US-09-252-991A-8200/C
Sequence 8200, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8200
LENGTH: 411
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8200

Query Match 1.6%; Score 17; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 CGGCAGTTGCCATCG 307
DB 342 CGGCAGTTGCCATCG 326

RESULT 18
US-09-621-976-2248
; Sequence 2248, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2248
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 129..446
US-09-621-976-2248

Query Match 1.6%; Score 17; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 CCAGCGTCCTCTCCCTC 147
Db 202 CCAGCGTCCTCTCCCTC 218

RESULT 19
US-09-252-991A-13127/c
; Sequence 13127, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13127
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13127

Query Match 1.6%; Score 17; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 TGGCGTCACCGCGGC 390
Db 457 TGGCGTCACCGCGGC 441

RESULT 20
US-08-797-812-25
; Sequence 25, Application US/08797812
; Patent No. 6228575
; GENERAL INFORMATION:
; APPLICANT: Gíngeras, Thomas A.
; APPLICANT: Mack, David
; APPLICANT: Chee, Mark S.

APPLICANT: Berro, Anthony J.
APPLICANT: Stryer, Lubert
APPLICANT: Chandour, Ghassan
APPLICANT: Wang, Ching
TITLE OF INVENTION: Chip-Based Species Identification and
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,812
FILING DATE: 07-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,765
FILING DATE: 15-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/629,031
FILING DATE: 08-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,631
FILING DATE: 01-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,339
FILING DATE: 08-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 16528X-018550
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 706 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-797-812-25

Query Match 1.6%; Score 17; DB 3; Length 706;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 GGGCACCGCATGAGT 343
Db 663 GGGCACCGCATGAGT 679

RESULT 21
US-09-252-991A-11363
; Sequence 11363, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11363
LENGTH: 843
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11363

Query Match 1.6%; Score 17; DB 4; Length 843;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 CCGCGCGAGATCCGCC 445
DB 783 CCGCGCGAGATCCGCC 799

RESULT 22
US-09-252-991A-11343
Sequence 11343, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11343
LENGTH: 930
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11343

Query Match 1.6%; Score 17; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 CCGCGCGAGATCCGCC 445
DB 811 CCGCGCGAGATCCGCC 827

RESULT 23
US-09-252-991A-2244/c
Sequence 2244, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2244
LENGTH: 984
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2244

Query Match 1.6%; Score 17; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 75;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 291 CGGCCAGTTCGCATCG 307
DB 572 CGGCCAGTTCGCATCG 556

RESULT 24
US-09-252-991A-11960/c
Sequence 11960, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11960
LENGTH: 993
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11960

Query Match 1.6%; Score 17; DB 4; Length 993;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 821 TGCACGGCTTGCCCTG 837
DB 588 TGCACGGCTTGCCCTG 572

RESULT 25
US-09-252-991A-2600
Sequence 2600, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2600
LENGTH: 999
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2600

Query Match 1.6%; Score 17; DB 4; Length 999;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 CGGCCAGTTCGCATCG 307
DB 723 CGGCCAGTTCGCATCG 739

RESULT 26
US-09-252-991A-8255
Sequence 8255, Application US/09252991A
Patent No. 6551795

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8255
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8255

Query Match          1.6%; Score 17; DB 4; Length 999;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      404 AGGAGTCGAGCGGAG 420
Db      278 AGGAGTCGAGCGGAG 294

RESULT 27
US-09-394-455-3
; Sequence 3, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustín, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPMR ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; FILE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USE
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/099,771
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)...(1008)
US-09-394-455-3

Query Match          1.6%; Score 17; DB 4; Length 1008;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      642 GACCTGAAGCGGAGAA 658
Db      454 GACCTGAAGCGGAGAA 470

RESULT 28
US-09-252-991A-12024
; Sequence 12024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11990
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11990

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12024
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12024

Query Match          1.6%; Score 17; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      821 TGCACGCGCTTGCGCTG 837
Db      245 TGCACGCGCTTGCGCTG 261

RESULT 29
US-09-252-991A-2312/c
; Sequence 2312, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2312
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2312

Query Match          1.6%; Score 17; DB 4; Length 1095;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      291 CGGCCAGTTCGCATCG 307
Db      331 CGGCCAGTTCGCATCG 315

RESULT 30
US-09-252-991A-11990/c
; Sequence 11990, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11990
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11990
```

Query Match 1.6%; Score 17; DB 4; Length 1119;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 821 TGACAGCGCTTGCGCTG 837
DB 740 TGACAGCGCTTGCGCTG 724

RESULT 31
US-09-252-991A-8104
Sequence 8104, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8104
LENGTH: 1134
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8104

Query Match 1.6%; Score 17; DB 4; Length 1134;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 CGGCGAGTCGCGCATCG 307
DB 388 CGGCGAGTCGCGCATCG 404

RESULT 32
US-09-107-532A-3059
Sequence 3059, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arianello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3059:
SEQUENCE CHARACTERISTICS:
LENGTH: 1281 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature
LOCATION: (1) LOCATION 1..1281
SEQUENCE DESCRIPTION: SEQ ID NO: 3059:
US-09-107-532A-3059

Query Match 1.6%; Score 17; DB 4; Length 1281;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 653 GGAGAACATCATGTTCG 669
DB 766 GGAGAACATCATGTTCG 782

RESULT 33
US-09-252-991A-12508
Sequence 12508, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 12508
LENGTH: 1290
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12508

Query Match 1.6%; Score 17; DB 4; Length 1290;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 TGCCGTCAGCGCGGC 390
DB 1079 TGCCGTCAGCGCGGC 1095

RESULT 34
US-09-023-655-293/c
Sequence 293, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 293:
SEQUENCE CHARACTERISTICS:
LENGTH: 1307 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNOT11
CLONE: 1344641
US-09-023-655-293

Query Match 1.6%; Score 17; DB 4; Length 1307;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1027 CACTGGGGCTGAGAGAT 1043
|||
Db 905 CACTGGGGCTGAGAGAT 889

RESULT 35
US-09-252-991A-8001
Sequence 8001, Application US/09252991A
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8001
LENGTH: 1314
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8001

Query Match 1.6%; Score 17; DB 4; Length 1314;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 291 CGGCGAGTTCGCATCG 307
|||

Db 636 CGGCGAGTTCGCATCG 652

RESULT 36
US-09-252-991A-2821/c
Sequence 2821, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2821
LENGTH: 1431
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2821

Query Match 1.6%; Score 17; DB 4; Length 1431;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 511 CTGTGTCCGGCGCGCA 527
|||
Db 633 CTGTGTCCGGCGCGCA 617

RESULT 37
US-09-620-312D-1058
Sequence 1058, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong U.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunhui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pf_fl_genes Version 1.0
SEQ ID NO 1058
LENGTH: 1450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS

LOCATION: (397)..(1176)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1450)
OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-1058

Query Match 1.6%; Score 17; DB 4; Length 1450;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 673 ACAAGCATGCAGCCAGC 689
Db 1201 ACAAGCATGCAGCCAGC 1217

RESULT 38
US-09-252-991A-8133/c
Sequence 8133, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8133
LENGTH: 1755
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8133

Query Match 1.6%; Score 17; DB 4; Length 1755;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 291 CGGCCAGTTCGCCATCG 307
Db 1666 CGGCCAGTTCGCCATCG 1650

RESULT 39
US-09-178-252-3
Sequence 3, Application US/09178252
Patent No. 6218188
GENERAL INFORMATION:
APPLICANT: Cardineau, Guy A.
APPLICANT: Stelman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: MA-714XC2
CURRENT APPLICATION NUMBER: US/09/178,252
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 60/065,215
EARLIER FILING DATE: 1997-11-12
EARLIER APPLICATION NUMBER: 60/076,445
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1815
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-178-252-3

Query Match 1.6%; Score 17; DB 3; Length 1815;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 840 CTGCTCAACATCCTG 856
Db 39 CTGCTCAACATCCTG 55

RESULT 40
US-09-178-252-7
Sequence 7, Application US/09178252
Patent No. 6218188
GENERAL INFORMATION:
APPLICANT: Cardineau, Guy A.
APPLICANT: Stelman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: MA-714XC2
CURRENT APPLICATION NUMBER: US/09/178,252
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 60/065,215
EARLIER FILING DATE: 1997-11-12
EARLIER APPLICATION NUMBER: 60/076,445
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 1815
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-178-252-7

Query Match 1.6%; Score 17; DB 3; Length 1815;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 840 CTGCTCAACATCCTG 856
Db 39 CTGCTCAACATCCTG 55

RESULT 41
US-09-178-252-8
Sequence 8, Application US/09178252
Patent No. 6218188
GENERAL INFORMATION:
APPLICANT: Cardineau, Guy A.
APPLICANT: Stelman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: MA-714XC2
CURRENT APPLICATION NUMBER: US/09/178,252
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 60/065,215
EARLIER FILING DATE: 1997-11-12
EARLIER APPLICATION NUMBER: 60/076,445
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1815
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-178-252-8

Query Match 1.6%; Score 17; DB 3; Length 1815;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 CTGCTCAACATCTCTG 856
|||||
Db 39 CTGCTCAACATCTCTG 55

RESULT 42

US-09-826-660-3
; Sequence 3, Application US/09826660
; Patent No. 6673990
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Steilman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-3

Query Match 1.6%; Score 17; DB 4; Length 1815;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 840 CTGCTCAACATCTCTG 856
|||||
Db 39 CTGCTCAACATCTCTG 55

RESULT 43
US-09-826-660-7
; Sequence 7, Application US/09826660
; Patent No. 6673990
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Steilman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-7

Query Match 1.6%; Score 17; DB 4; Length 1815;
Best Local Similarity 100.0%; Pred. No. 76;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 840 CTGCTCAACATCTCTG 856
|||||
Db 39 CTGCTCAACATCTCTG 55

RESULT 44

US-09-826-660-8
; Sequence 8, Application US/09826660
; Patent No. 6673990
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Steilman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-8

Query Match 1.6%; Score 17; DB 4; Length 1815;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 840 CTGCTCAACATCTCTG 856
|||||
Db 39 CTGCTCAACATCTCTG 55

RESULT 45
US-09-336-643A-11
; Sequence 11, Application US/09336643A
; Patent No. 6399761
; GENERAL INFORMATION:
; APPLICANT: Miller, Andrew P.
; APPLICANT: Curran, Mark Edward
; APPLICANT: Hu, Ping
; APPLICANT: Rutey, Marc
; APPLICANT: Wang, Jian-Wang
; TITLE OF INVENTION: No. 6399761 Human Potassium Channels
; FILE REFERENCE: SEQ-15P
; CURRENT APPLICATION NUMBER: US/09/336,643A
; CURRENT FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 60/076,687
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 60/116,448
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/US99/03826
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: H. sapiens
; NAME/KEY: CDS
; LOCATION: (383)...(1157)

OTHER INFORMATION: K+Hnov15
US-09-336-643A-11

Query Match 1.6%; Score 17; DB 4; Length 1862;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 673 ACAAGCATGCAGCCAGC 689
Db 1187 ACAAGCATGCAGCCAGC 1203

Search completed: June 30, 2004, 21:26:53
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 30, 2004, 19:02:46 ; Search time 3172 Seconds

(without alignments)
3398.565 Million cell updates/sec

Title: FIGURE9
Perfect score: 1826
Sequence: 1 MXMIGLIGRSSPXQJDSLS.....GXEMSLIGLIDMPGPHQXRTX 361

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgnt1/uspt0/spool/GBRBYESUS41/runatc_29062004_143151_24567/app_query.fasta_1.519
-DB=EST -QMT=fastap -SUFFIX=xsc -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=exc -HEAPSIZE=550 -MINLEN=0 -MAXLEN=2000000000
-USER=GBRBYESUS41.qcgn1_1_2810_@runatc_29062004_143151_24567 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_estcum:*
3: em_estlin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	738	40.4	666	13	BU670814
2	730.5	40.0	947	13	BU841154
3	725	39.7	597	13	BU671103
4	725	39.7	971	13	BU513845
5	722	39.5	844	14	CF617929
6	722	39.5	850	14	CF617168
7	722	39.5	972	12	B1554559
8	722	39.5	976	11	BC012210
9	714.5	39.1	681	12	B149425
10	710	38.9	637	14	CA334157
11	708.5	38.8	887	12	EG968191
12	699.5	38.3	936	13	BQ945722
13	689	38.3	966	10	BF012103
14	695.5	38.1	710	12	BG277312
15	690	37.8	703	12	BG865492
16	688.5	37.7	1204	11	BC023238
17	683.5	37.4	756	10	BF162798
18	677	37.1	584	9	AA763351
19	670	36.7	600	14	CA529114
20	666.5	36.5	723	12	BG968192
21	666.5	36.5	743	12	B1690902
22	666	36.5	871	12	B1652999
23	664	36.4	785	12	B1904738
24	657.5	36.0	707	12	B1904738
25	654.5	35.8	622	12	B1417443
26	653	35.8	663	10	BF019568
27	638.5	35.0	1039	12	BG419640
28	638	34.9	997	13	BC071271
29	634.5	34.7	671	10	BF19588
30	630.5	34.5	961	10	BF19588
31	629	34.4	557	9	AA12667
32	628	34.4	553	10	AA7442180
33	624.5	34.2	1127	10	BE733365
34	624	34.2	708	10	AA746323
35	624	34.2	731	10	BE306065
36	621	34.0	1022	13	BQ069289
37	620.5	34.0	1058	12	BG421064
38	618	33.8	1037	13	BQ071749
39	616.5	33.8	798	12	BG862263
40	606	33.2	565	10	BF547702
41	600.5	32.9	729	12	BM929308
42	597	32.7	465	9	AA017803
43	590	32.3	455	14	CB741392
44	588.5	32.2	855	12	BF828782
45	580	31.8	796	12	BG421646

ALIGNMENTS

RESULT 1
BU670814
LOCUS
DEFINITION NISC_1r01e11.y1 NCI CGAP Pr49 Rattus norvegicus cDNA clone
IMAGE:5598044.5', mRNA sequence.
ACCESSION BU670814
VERSION BU670814.1 GI:23398790
KEYWORDS EST
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 666)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
info@image.lnl.gov
Plate: L1AM12384 row: 1 column: 21
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..666
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:5598044"
/sex="male"
/tissue_type="ventral prostate, pool of 3-, 5-, and 7-days
post-castration"
/dev_stage="adult, 11 week"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Pr49"
/note="Organ: prostate; Vector: PCMV-SPORT6.1; Site: 1:
NotI; Site 2: EcoRV; Cloned unidirectionally. Primer:
Oligo dt. Pool of 3 primary libraries: NCI_CGAP_Pr30
(ventral prostate from 11 wk male, 3 days
post-castration, average insert size 2 kb), NCI_CGAP_Pr40
(ventral prostate from 11 wk male, 5 days
post-castration, average insert size 1.6 kb) and
NCI CGAP Pr41 (ventral prostate from 11 wk male, 7 days
post-castration, average insert size 2.5 kb). Constructed
by life technologies/Invitrogen. Note: this is a NCI_CGAP
Library."

ORIGIN

Alignment Scores:

Pred. No.:	3,27e-41	Length:	666
Score:	738.00	Matches:	163
Percent Similarity:	80.37%	Conservative:	9
Best Local Similarity:	76.17%	Mismatches:	25
Query Match:	40.42%	Indels:	17
DB:	13	Gaps:	4

FIGURE9 (1-361) x BU670814 (1-666)

QY 43 ArgAlaserValLeuSerLeuValAlaIleProlycCysLeuSer**GlySerLeuGly 62
Db 67 CGAAGGGCATGGCTGCTCCCTGCTGCTCCGAGATCTCGAAGAGACAGCGCCATG 126
QY 63 SerSerValValValGlyAsnLeuGlyThrAspAlaGlnArgLeuGlyTyrArgValLeu 82
Db 127 TCCACGTTCCAGGCGAGGAGAGCTT--GAGGACCATTTATGATGAGAGGAGCTT--- 180
QY 83 ProSerSerGlySerAlaAlaLeuSerCysSerPheProHisSerGlyPheAlaIleVal 102
Db 181 ---GGCAGCGCGCGAG-----TTCCGCCATCGCG 204
QY 103 ArgGlyCys-----LysGlyThrGlyMetGlnTyrAlaAlaLysPheIleLysArg 120
Db 205 CGCAGTGCACGACGAGGAGGCGGATGAGTACCGGCGCAGTTCAAAAGAGCGG 264
QY 121 ArgLeuProSerSerArgArgGlyValSerArgGlnGlnIleGlnArgGlnValSerIle 140
Db 265 CGCTGCGCTCCAGCGCGCGGCTGAGCGCTGAGAGATCGAGGCGCGAGCACTC 324
QY 141 LeuArgGlnIleArgHisProAsnIleIleThrLeuHisAspValPheGluAsnLysTyr 160

Db 325 CTGGCGGAGATCCGCCACCCCAACATCATCATCGCTGCACGATGTCTTGAGACAGACA 384
QY 161 AspValValLeuIleLeuGlnLeuValSerGlyGlyGlnLeuPheAspPheLeuAlaGln 180
Db 385 GATGTGAGTCTATCTTGAGAGCTGTGTCCGCGGCGAACTTTTCGACTTTCTGCTGAG 444
QY 181 LysAsp-His**GlnArgMetArgProArgSerSerSerSerArgSerTyrPheValSer 200
Db 445 AGGAGTCACTGACAGAGATGAGCGCAGCGAGTTCTTCAAGACATCTTGACCGGTTC 504
QY 200 rThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgArgThrSerCysCy 220
Db 505 CACTACTGTCATCTCCAGAGGATCGCGCATTTGACCTGAAGCCGAGACATCATGTGTG 564
QY 220 rThrThrSerMetGlnPheAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGln 240
Db 565 CTGACAGACATGACGACCGACCGCCAGCATTAAGCTTCATGACTTTGGCATGCGCAGG 624
QY 240 YSerArgProValAlaSerSerArgThrSerLeuAlaArg 253
Db 625 ATCGAGGCGCGTAGCGAGTTCAAGACATCTTTGSCACGC 664

RESULT 2
BU641154 947 bp mRNA linear EST 16-OCT-2002
LOCUS
DEFINITION
AGENCOURT 10187633 NIH MGC 134 Mus musculus cDNA clone
IMAGE:6518980 5', mRNA sequence.
BU641154
ACCESSION
BU641154.1 GI:24025554
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 947)
Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
info@image.lnl.gov
Plate: L1AM14101 row: 1 column: 05
High quality sequence stop: 456.
Location/Qualifiers
1..947
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6518980"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_134"
/note="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by Resgen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.85e-40	Length:	947
Score:	730.50	Matches:	167
Percent Similarity:	71.49%	Conservative:	11
Best Local Similarity:	67.07%	Mismatches:	33
Query Match:	40.01%	Indels:	38
DB:	13	Gaps:	3

FIGURE9 (1-361) x BU641154 (1-947)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 176 TTGGCCATCGTGGCAGATGCGCAGCAGAGGAGCAGCGGATGAGATGAGCGCAAGTTTC 235
QY 117 IleLysIleValArgLysProSerSerArgArgLysValSerArgLysGluIleGluArg 136
Db 236 ATCAAGAGGCGGCGCTGCTCCATCCAGCGCGGCTGCTGAGCCGAGAGAGATCGAACGC 295
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 296 GAGGTGAGCATCTCGCGCGAGATCCGCAACCATCATATACATGATGATGATGATGATGAT 355
QY 157 GluAsnLysThrAspValValIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 356 GAGAACCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 415
QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSerSerArgS 196
Db 416 TTCTGCGCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 475
QY 196 rTTPThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db 476 CTAGACGGTTCACATCCGACCTCCAGCGCATCGACATCTTACCTGAGCCGAG 535
QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 536 AACATCATGTTGCTGAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 595
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSer 256
Db 596 ATCGCGCAGACAGATCGAGCTGCGCAGAGATTCAGAACATCTTTGGCA----- 644
QY 256 rSerValArgGlyArgGlyGlyHisHisProIleGly**IleLeuHisGlyLeuGly 276
Db 644 ----- 644
QY 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThr 296
Db 645 -----CACCAGATTGTCGCCCCCCCNAGATCGAGATGAGATGAGATGAGATGAGAT 691
QY 296 rLeuGlySer**TyrValGluHisArgArgHisHisLeuHisPro-ValSerAla*** 315
Db 692 GGCCTGAGAGCTGACCTGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 745
QY 316 AspGlyGlnGlyProGlnThrValProAlaArgGlyProGlyIleArgAlaGlyThr 335
Db 746 -----GAGAGCTCCCATCTTCGCGGCGAGAGACAGCAGAGACAGCAGAGACAGC 796
QY 336 AlaAsnCysLysHisTrpGly 342
Db 797 ATCTCCGAGTGAACATATGGA 817

```

CDNA Library Preparation:
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/MLNL
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLNL at:
 info@image.llnl.gov
 Plate: LLM12385 row: C column: 7
 Seq primer: M13R1 reverse primer (ABI).
 Location/Qualifiers
 1..597
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:5598270"
 /sex="male"
 /tissue_type="ventral prostate, pool of 3-, 5-, and 7-days
 post-castration"
 /dev_stage="adult, 11 week"
 /lab_host="VDH103 (T1 phage-resistant)"
 /clone_id="NCI CGAP Pr49"
 /note="Organ: prostate; Vector: pCMV-Sport6.1; Site 1:
 NotI; Site 2: EcoRV; Cloned unidirectionally. Primer:
 Oligo dt. Pool of 3 primary libraries: NCI CGAP Pr30
 (ventral prostate from 11 wk male, 3 days
 post-castration, average insert size 2 kb), NCI CGAP Pr40
 (ventral prostate from 11 wk male, 5 days
 post-castration, average insert size 1.6 kb) and
 NCI CGAP Pr41 (ventral prostate from 11 wk male, 7 days
 post-castration, average insert size 2.5 kb). Constructed
 by life technologies/Invitrogen. Note: this is a NCI CGAP
 Library."

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Gaps:
2,196-40	597	725.00	95.62%	95.00%	39.70%	13	1

FIGURE9 (1-361) x BU671103 (1-597)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 120 TTGGCCATCGTGGCAGATGCGCAGCAGAGGAGCAGCGGATGAGATGAGCGCAAGTTTC 179
QY 117 IleLysIleValArgLysProSerSerArgArgLysValSerArgLysGluIleGluArg 136
Db 180 ATCAAGAGGCGGCGCTGCTCCATCCAGCGCGGCTGCTGAGCCGAGAGATCGAACGC 239
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 240 GAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCATCAGCTGACAGATGTTTC 299
QY 157 GluAsnLysThrAspValValIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 300 GAGAACCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSer-SerArgS 196
Db 360 TTCTGCGCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419
QY 196 rTTPThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db 420 CTGAGCGGTTCACATCCGACCTCCAGAGGATCCCGGACATTTGACCTGAGACCCGGA 479
QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 480 AACATCATGTTGCTGAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 539
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGln 254

```

```

Db      540 CATCGCGCAGAGATCGAGCGCGGAGGAGTTCACAGACATCTTGGCAGCGCAG 595
RESULT 4
BUI51845
LOCUS   BUI51845                971 bp    mRNA    linear    EST 12-SEP-2002
DEFINITION AGENCOURT_10110388 NIH_MGC_134 Mus musculus cDNA clone
IMAGE:6510429 5', mRNA sequence.
ACCESSION BUI51845
VERSION   BUI51845.1 GI:22821371
KEYWORDS EST.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 971)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cga@bbs-remail.nih.gov
          Tissue Procurement: Dr. David Rowe
          cDNA Library Preparation: Invitrogen Corp
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LLM14079 row: e column: 22
          High quality sequence stop: 545.
          Location/Qualifiers
            1..971
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /clone="IMAGE:6510429"
              /tissue_type="undifferentiated limb"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_134"
              /note="Vector: PCMV-SPOB-1; Site 1: EcoRV; Site 2: NotI;
              Cloned unidirectionally. Primer: Oligo dT. Average insert
              size 1.7 kb. Constructed by Resgen, Invitrogen Corp. Note:
              this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      4,61e-40      Length:      971
Score:          725.00      Matches:      159
Percent Similarity: 77.31%      Conservative: 8
Best Local Similarity: 73.61%      Mismatches: 18
Query Match:    39.70%      Indels:      32
DB:             13          Gaps:          2

FIGURE9 (1-361) x BUI51845 (1-971)
QY      99 PheAlAlIeVAlAgLyS Cys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db      21 TTTCGCATCGCGCCAGATGCGCAGAGAGGCGCGGCGATGAGATGACGCGCAGATTC 80
QY      117 IleLysLysArGArGluProSerSerArGArGlyValSerArGluGluIleGluArg 136
Db      81 ATCAAGAGCGCGCCCTGCGCATCCAGCGCGCGCGGTGTAGCCCGGAGGAGATCGAAGC 140
QY      137 GluValSerIleLeuArGluIleArgIleProAsnIleIleThrIleuHisAspValPhe 156
Db      141 GAGGTGAGCATCCGCGCCGAGATCCGCGCACCCCACTAACAATGATGACATGACGTTC 200
QY      157 GluAsnLysThrAspValValIleuIleGluIleuValSerGlyGlyGluIleuPheAsp 176
Db      201 GAGAAACAAGACAGATGAGTGTGATCTGAGCTGTGTCCGTTGCGGAGCTTTTCGAC 260
QY      177 PheLeuAlaGluLysAsp-His**GlnArGMeLarProArGSerSerSerArGSe 196
          |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:

```

```

Db      261 TTCTGCGCGAGAGAGACTATTGACGAGGATGAGCCAGCATCTCTCAACAAATC 320
QY      196 rThrThrValSerThrThrCysThrProSerAlaSerArgThrIleuThr**SerArgAr 216
Db      321 CTAGACGGTGTCCATCTACTCTGACATCCAGGCGATCGACACTTTGACTGAGCCGAG 380
QY      216 gThrSerCysCysThrThrSerMetGlnProAlaHisAlaLeuSerSerThrIleuAl 236
Db      381 AACATCATGTGCTCTGACAGACGACGACGACGACGACGACGACGACGACGACGACG 440
QY      236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerIleuAlaArgInSerSe 256
Db      441 ATCGCGCACAGATGAGGCTGCGACGCGCATTCAGAAACATCTTGGCA----- 489
QY      256 rSerValArgGlyArgCysGlyHisHisProIleGly**IleuHisGlyLeuGlyLe 276
Db      489 -----
QY      276 uThrCysLeuAsnAspProValPheHisSerPro**AspCysLysLeu**ThrThrTr 296
Db      490 -----CACCGGAG-TTGTGCGCCCGCCGACATGTGACTATGAGCCACTTG 535
QY      296 PheuGlySer**TyrValGluHisArgArgHisHisIleuHisPro 311
Db      536 GCTTGAGGCTGACATGTGAGACATTGGCGTCATCATCATCATCTCT 581
RESULT 5
CF617929                844 bp    mRNA    linear    EST 01-OCT-2003
LOCUS   CF617929
DEFINITION AGENCOURT_15765794 NIH_MGC_222 Mus musculus cDNA clone
IMAGE:30535626 5', mRNA sequence.
ACCESSION CF617929
VERSION   CF617929.1 GI:37235384
KEYWORDS EST.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 844)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT   Contact: Daniela S. Gethard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10x07 Bethesda, MD 20892
          Email: cga@bbs-remail.nih.gov
          Tissue Procurement: Naryan Bhat
          cDNA Library Preparation: Express Genomics
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: NDM607 row: b column: 03
          High quality sequence stop: 668.
          Location/Qualifiers
            1..844
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /clone="IMAGE:30535626"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_222"
              /note="Organ: Placenta; Vector: pExpress-1; Site 1: EcoRV;
              Site 2: NotI; RNA obtained from three placentas from
              female C57/BL6 mouse at 16 days pregnancy. Tissues were
              snap-frozen and kept at -80C for two days before RNA
              extraction and purification (TRI-reagent method). cDNA was
              primed using oligo-dT primer:
              5'-DCACTAGTCTAGATCGGAGCGGCGCCCT(T)25-3' and cloned into
              the EcoRV/NotI sites of pExpress-1. Size-selection >1 kb
              resulted in an average insert size of 1.5 kb. Library is

```



```

Db      455 CTAGACGGTGTCCACTCTGCACTCCAGGCAATCCACACTTTGACCTGAAGCCCGAG 514
Qy      216 gThSerCysCSTPThrSerMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236
Db      515 AACATCATGTTGCTGGACAAACAGCAGCCGCCCGCATTAAGCTCATGACTTTGGC 574
Qy      236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInserSe 256
Db      575 ATCCGGCAGACGATCGAGGCTGGCGAGTTCAGAAACATCTTTGCCA----- 623
Qy      256 rSerValArgGlyArgCysGlyHisPro-IleGly***IleLeuHisGlyLeuGlyL 276
Db      624 -----CACCCGAGTT----- 633
Qy      276 eutThrCysLeuAsnAsnProValPheHisSerPro**AspCysLeuLeu**ThrThrT 296
Db      634 -----TGT-----CGCCCCGAGATCGCAACTGACCTATGAGCCACTT 667
Qy      296 rPheLeuGlySer***TyrValGlnHisArgArgHisHisLeuHisProValSerAla**A 316
Db      668 GGCTTGGAGGCTGACATGTGGAGCATTTGGCGCTCATCCCTACATCTCTCTGAG----- 720
Qy      316 spGlyGlnGlyProGlnThrValProAlaArgGly 327
Db      721 -----CGAGAGCTCCCATTCCTGGG 741

RESULT 7
B1554559 972 bp mRNA linear EST 05-SEP-2001
LOCUS 603235693P1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5310046 5'
ACCESSION B1554559
VERSION B1554559.1 GI:15441873
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga9bs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L19M11785 row: 1 column: 23
High quality sequence stop: 845.
Location/Qualifiers
1..972
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5310046"
/lab_host="MDH10B (T1 phase-resistant)"
/clone_lib="NCI_CGAP_L19"
/clone="Organ: liver; Vector: pCMV-SPORTS; Site: 1: NCI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP library."
ORIGIN
Alignment Scores: 7.44e-40 Length: 972
Pred. No.: 722.00 Matches: 163
Score: Percent Similarity: 73.82% Conservative: 9

```

```

Best Local Similarity: 69.96% Mismatches: 21
Query Match: 39.54% Indels: 40
DB: 12 Gaps: 5
FIGURE9 (1-361) x B1554559 (1-972)
Qy      99 PheAlaIleValArgLeuCys-----TysGlyThrGlyMetGlyTyrAlaAlaPhe 116
Db      65 TTGGCCATGTCGCAAGTGCAGAGCCAGAGGCGACGGGCAATGAGTATGACGCAAGTTC 124
Qy      117 IleLysLysArgArgLeuProSerSerArgArgGlyValaSerArgGlnGlnIleGlnArg 136
Db      125 ATCAAGAGAGCGCGCGCTGCATCCAGCCGCGGGTGTGAGACCGGAGAGATGCAAGCGC 184
Qy      137 GluValSerIleLeuArgGlnIleArgHisProAlaIleIleThrLeuHisAspValPhe 156
Db      185 GAGGTGAGCATCTCTGGAGATCGCCACCCCAACATCAATGACATGACGCTTTC 244
Qy      157 GluAsnLysThrAspValValLeuIleLeuGlnLeuValSerGlyGlyLeuPheAsp 176
Db      245 GAGAACAGACAGATGCTGTGCTGATCCTGAGAGCTGATCGGTGGGAGACTTTTGAC 304
Qy      177 PheLeuAlaGlnLysAsp-His***GlnArgYechArgProArgSerSerSerArgSe 196
Db      305 TTCCTGGCCGAGAGAGATGATTCAGAGAGATGAGCGACAGTTCTCAACAAATC 364
Qy      196 rThrThrValaSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db      365 CTAGACGGTGTCCACTCTGCACTCCAGGCAATCCGACACTTTGACTGAAGCCCGAG 424
Qy      216 gThSerCysCSTPThrSerMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236
Db      425 AACATCATGTTGCTGGACAAACAGCAGCCGCCCGCATTAAGCTCATGACTTTGGC 484
Qy      236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInserSe 256
Db      485 ATCCGGCAGACGATGAGAGCTGGCAGCGATTCAGAAACATCTTTGCA----- 533
Qy      256 rSerValArgGlyArgCysGlyHisPro-IleGly***IleLeuHisGlyLeuGlyL 276
Db      534 -----CACCCGAGTT----- 543
Qy      276 eutThrCysLeuAsnAsnProValPheHisSerPro**AspCysLeuLeu**ThrThrT 296
Db      544 -----TGT-----CGCCCCGAGATCGCAACTGACCTATGAGCCACTT 577
Qy      296 rPheLeuGlySer***TyrValGlnHisArgArgHisHisLeuHisProValSerAla**A 316
Db      578 GGCTTGGAGGCTGACATGTGGAGCATTTGGCGCTCATCCCTACATCTCTCTGAG----- 630
Qy      316 spGlyGlnGlyProGlnThrValProAlaArgGly 327
Db      631 -----CGAGAGCTCCCATTCCTGGG 651

RESULT 8
BC012210 976 bp mRNA linear HTC 19-NOV-2003
LOCUS BC012210
DEFINITION Mus musculus death-associated kinase 3, mRNA (cDNA clone
IMAGE:3993055), with apparent retained intron.
ACCESSION BC012210
VERSION BC012210.1 GI:15126556
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Title: Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Wax, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

```

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggiano, N.A., Peters, G.J., Abramson, R.D., Muliyil, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.B., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

12477932

2 (bases 1 to 976)

Strausberg, R.

Direct Submission

Submitted (06-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hui, S.W., Loussag, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>

Series: IRAX Plate: 23 Row: n Column: 16

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6681132

This clone has the following problem: retained intron.

Location/Qualifiers

1. 976

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:3993055"

/tissue_type="Mammary tumor. MAP-TGF alpha model. 7 months old. Gross tissue."

/clone_id="NCI CGAP_Mam5"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 7.49e-40 Length: 976

Score: 722.00 Matches: 163

Percent Similarity: 73.82% Conservative: 9

Best Local Similarity: 69.96% Mismatches: 21

Query Match: 39.54% Indels: 40

DB: 11 Gaps: 5

FIGURE9 (1-361) x BC012210 (1-976)

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGlyIrrAlaAlaLysPhe 116

Db 164 TTTCGATCGTGGCCAGATGCGACAGAGGCGACGGGCAATGAGTATGACGCAAGTTTC 223

QY 117 IleLysIleValArgLysProSerSerArgArgGlyValSerArgGluGluLeuArg 136

Db 224 ATCAAGAGAGCGCGCGCTCGATCCAGCGCGCGGTGTGAGCGGAGAGATCGAAGCC 283

QY 137 GluValSerIleLeuArgGluLeuArgHisProAlaIleIleThrLeuHisAspValPhe 156

Db 284 GAGGTGAGATCTCTCGGAGATCGGACCCCAACATCATTAACCTCATGACGTGTTG 343

QY 157 GluAlaIleThrAspValValLeuIleLeuGluLeuValSerGlyGluLeuPheAsp 176

Db 344 GAGAACAGACAGATGTGTGTCTGATCTGAGCTGTGTCTCGGAGAGCTTTTCAGC 403

QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSerArgSe 196

Db 404 TTCCCTGCGCGAAGAGATGATTCAGAGAGATGAGGCGACAGTTCCTCAACAAATC 463

QY 196 rTrpThrValSerTrpThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216

Db 464 CTAGAGGTTCTCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 523

QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236

Db 524 AACATCATGTTCTGTGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 583

QY 236 aSerArgThrArgLysArgProValAlaSerSerArgThrSerLeuAlaArgLysSerSe 256

Db 584 ATCGGCGACAGATGAGAGCTGGACGAGATTCAGAAATCTTTGGCA----- 632

QY 256 rSerValArgGlyArgCysGlyHisHisPro-IleGly***IleLeuHisGlyLeuGlyL 276

Db 633 -----CACCGAGTT----- 642

QY 276 eutThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrT 296

Db 643 -----TGT-----CGCCCGGAGATCGTGAACATGAGGACCACTT 676

QY 296 rPLeuGlySer***TyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316

Db 677 GCGTTGAGAGCTGACATGTGAGCATTTGGCGTCATCATCTCTCTCTGAG----- 729

QY 316 spGlyGlnGlyProGlnThrValProValArgGly 327

Db 730 -----CGAGCGGCCCATTCCTGGG 750

RESULT 9

BI149425 681 bp mRNA linear EST 05-JUL-2001

LOCUS 602848986F1 NCI CGAP Lu29 Mus musculus cDNA IMAGE:5012517 5',

DEFINITION mRNA sequence.

ACCESSION BI149425

VERSION BI149425.1 GI:14609426

KEYWORDS EST

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

1 (bases 1 to 681)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>

Plate: LLM11063 row: 1 column: 22

High quality sequence stop: 681.

Location/Qualifiers

1. 681

FEATURES

source

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:5012517"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu29"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: Salt;
Site 2: NCI; Cloned unidirectionally. Primer: oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

ORIGIN

Alignment Scores:

Pred. No.:	1,43e-39	Length:	681
Score:	714.50	Matches:	159
Percent Similarity:	76.50%	Conservative:	7
Best Local Similarity:	73.27%	Mismatches:	18
Query Match:	39.13%	Indels:	33
DB:	12	Gaps:	4

FIGURE9 (1-361) x B1149425 (1-681)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 113 TTTCGCATCGCGCCAGTGCAGAGAGGCGCGGCGATGAGTATGACCGCCAGTTTC 172
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 173 ATCAGAGAGCGCGCGCTCCATCCAGCGCGCGGTGTGAGCCGGAGAGATCGAAGCG 232
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 233 GAGGTGAGCATCCCGCGAGATCCGCCACCCCAACATCAACATGACATGACGCTTC 292
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 293 GAGAACAGACAGATGAGTGTCTGATCTGAGCTGTGCTGCGGTGCGAGCTTTTCGAC 352
QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSerArgSer 196
DB 353 TTCCGGCGCGAGAGAGATGACGAGAGATGAGCGACGAGTTCTCTCAACAGATC 412
QY 196 rTTPThrValSerThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 413 CTAGCGGTGTCCTACTGCTGCTCCAGCGCATGCGACCTTGAACCTGAAAGCCCGAC 472
QY 216 gThrSerCysCSTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 473 AACATCATGTTGCTGAGCAAGCAGCGCGCCAGCCCGCATTAAGCTCATCGACTTTGGC 532
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInserSe 256
DB 533 ATCGCGCAGAGATCGAGCTGCGAGAGTTCAAGAAACATCTTTGGCA----- 581
QY 256 rSerValArgGlyArgCysGlyHisPro-IleGly***IleLeuHisGlyLeuGlyL 276
DB 582 -----CACCGAGTT----- 591
QY 276 eutThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu***ThrThr 296
DB 592 -----TGT-----CGCCCCCGAGATCGTGAACCTATGAGCCACTT 625
QY 296 rPLeuGlySer***TyrValGluHisArgArgHisHisLeuHisPro 311
DB 626 GGCTTGGAGGCTGACATGTGAGCAATTGGCGCTATCCCTACCTACCTCT 672

```

RESULT 10 CA334157 637 bp mRNA linear EST 04-NOV-2002
 LOCUS CA334157
 DEFINITION NISC_1606b10.y1 NCI_CGAP_Pr50 Rattus norvegicus cDNA clone

IMAGE:5599435 5', mRNA sequence.
 CA334157
 VERSION CA334157.1 GI:24552255
 KEYWORDS EST
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 637)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 CDNA Library Preparation:
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
 Plate: LLM12388 row: C column: 20
 Seq primer: M13RP1 reverse primer (ABI).

FEATURES

source
 1. 637
 location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:5599435"
 /sex="male"
 /tissue_type="dorsolateral prostate, pool of 3-, 5-, and
 7-days post-castration"
 /dev_stage="adult, 11 week"
 /lab_host="DH10B (TI phase-resistant)"
 /clone_lib="NCI CGAP_Pr50"
 /note="Organ: prostate; Vector: pCMV-SPORT6.1; Site 1:
 NotI; Site 2: EcoRV; Cloned unidirectionally. Primer:
 oligo dr. Pool of 3 primary libraries: NCI CGAP_Pr39
 (dorsolateral prostate from 11 wk male, 3 days
 post-castration, average insert size 2.7 kb),
 NCI CGAP_Pr29 (dorsolateral prostate from 11 wk male, 5
 days post-castration, average insert size 2.2 kb) and
 NCI CGAP_Pr42 (dorsolateral prostate from 11 wk male, 7
 days post-castration, average insert size 2.2 kb).
 Constructed by Life Technologies/Invitrogen. Note: this is
 a NCI_CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.64e-39	Length:	637
Score:	710.00 <td>Matches:</td> <td>146 </td>	Matches:	146
Percent Similarity:	96.71% <td>Conservative:</td> <td>2 </td>	Conservative:	2
Best Local Similarity:	96.05% <td>Mismatches:</td> <td>3 </td>	Mismatches:	3
Query Match:	38.88% <td>Indels:</td> <td>1 </td>	Indels:	1
DB:	14	Gaps:	1

FIGURE9 (1-361) x CA334157 (1-637)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 184 TTTCGCATCGTGCAGAGTGCAGAGAGGCGCGGCGATGAGTATGACCGCCAGTTTC 243
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 244 ATTAAGAGAGCGCGCGCTCCGTCAGCGCGCGGTGTGAGCCCGTGAAGATCGAGCGC 303
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 304 GAGGTGAGCATCTCCGAGAGTCCGCCACCCCAACATCAACATGACGCTGACAGATGTTC 363
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176

```

```

Db      164 GAGAACAGACAGATGCTGATCTTGGAGCTGTGTCGGCGGCGAACTTTTCAGC 423
Qy      177 Pheulnlaagluysasp-His***GlnArgMetArgProArgSerSerSerArgse 196
Db      424 TTCTGGCTGAGAGAGCTCAGACAGAGATGAGGCGACGAGTTCCTCAAGCAGATC 483
Qy      196 rTTPThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db      484 CTGACGCGTGTCCACTCTGCACTCCAGCGCATGCGCATTTGACCTGGAAGCGGAG 543
Qy      216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236
Db      544 AACATCATGTTGCTGAGCAGCATGACGCGCCAGCATTAAGCTCATGACTTTGGC 603
Qy      236 aSerArgThrGlySerArgProValAlaSerSer 247
Db      604 ATCGCGCACAGATCGAGCGCGTAGGACTTCA 637

RESULT 11
Bg968191 887 bp mRNA linear EST 12-JUN-2001
LOCUS 60283567F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4990027 5',
DEFINITION mRNA sequence.
Bg968191
ACCESSION Bg968191.1 GI:14355828
VERSION Bg968191.1
KEYWORDS Mus musculus (house mouse)
SOURCE EST.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 887)
REFERENCE NTH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L1AM1005 row: c column: 20
High quality sequence stop: 820.
FEATURES
Source
Location/Qualifiers
1..887
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4990027"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1b="NCI_CGAP_Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NciI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Alignment Scores:
Pred. No.: 5.56e-39 Length: 887
Score: 708.50 Matches: 161
Percent Similarity: 74.78% Conservative: 8
Best Local Similarity: 71.24% Mismatches: 22
Query Match: 38.80% Indels: 35
DB: 12 Gaps: 5
FIGURE9 (1-361) x Bg968191 (1-887)
Qy 99 PhealalleValArgLysCys-----LysGlyThrGlyMetGlyLysTrpAlaAlaLysPhe 116
|||||

```

```

Db      164 TTTCATCTGTCGCGAAGTGCAGAGAGGCGACGGGCGATGAGTATGACGCCAAGTTC 223
Qy      117 lLeuLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluileGluArg 136
Db      224 ATCAAGAAAGCGGCGCTCGATCCAGCCGCGGCTGTGAGCGGAGAGAGATCCAAAGCC 283
Qy      137 GluValSerileLeuArgGluileArgHisProAsnileileThrLeuHisAspValPhe 156
Db      284 GAGGTGAGATCTCTCGCGAGATCCGCCACCCCAACATCATTAACCTCATGACGTGTC 343
Qy      157 GluAsnLysThrAspValValLeuileLeuGluileuValSerGlyGlyGluLeuPheSp 176
Db      344 GAGAACAGACAGATGTGTGTCTGATCTTGAGGCTGTGTCGGTGGGAGACTTTTCAC 403
Qy      177 Pheulnlaagluysasp-His***GlnArgMetArgProArgSerSerSerArgse 196
Db      404 TTCTGGCGGAGAAAGAGATCATTTGACGAGATGAGCGCACGACTTCTCTCAACAATC 463
Qy      196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db      464 CTAGACGGTGTCCACTACTCTGCACTCCAGCGCATCGCACACTTGAACCTGAAAGCCGAG 523
Qy      216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236
Db      524 AACATCATGTTGCTGAGCAGCATGACGCGCCAGCATTAAGCTCATGACTTTGGC 583
Qy      236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db      584 ATCGCGCACAGATGAGAGCTGGCGAGCATTCAGAAACATCTTTGCA----- 632
Qy      256 rSerValArgGlyArgCysGlyHisHisPro-IleGly***IleLeuHisGlyLeuGlyL 276
Db      633 -----CACCCGAGTT----- 642
Qy      276 euThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
Db      643 -----TGT-----CGCCCCGAGACGTGAATTAAGACGACTT 676
Qy      296 rPLeuGlySer***TyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316
Db      677 GCTTGGAGGCTGACATGTGAGCATTTGCGTGCATCACTACAT-----TCCTCTGAA 730
Qy      316 sPGLyGlnGlyPro 320
Db      731 GCGGAGCGTCCCA 744

RESULT 12
Bg945722 936 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT 8953288 NCI CGAP_Co24 Mus musculus cDNA clone
DEFINITION IMAGE:6477473 5', mRNA sequence.
Bg945722
ACCESSION Bg945722.1 GI:22361200
VERSION Bg945722.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 936)
REFERENCE NTH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L1AM14019 row: h column: 18
High quality sequence stop: 708.

```

FEATURES
source

Location/Qualifiers
1. .936
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6477473"
/lab_host="DH10B (TI phage-resistant)"
/clone_1ib="NCI CGAP Co24"
/note="Organ: colon; Vector: PCWV-SPORT6; Site: 1; NctI; Site 2: SalI; cloned unidirectionally. Primer: oligo dt. Average insert size 1.6 Kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.: 2,536-38 Length: 936
Score: 699.50 Matches: 168
Percent Similarity: 64.73% Conservative: 10
Best Local Similarity: 61.09% Mismatches: 41
Query Match: 38.31% Indels: 57
DB: 13 Gaps: 5

FIGURE9 (1-361) x BQ945722 (1-936)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 110 TTTCCTCATCGCGCCAGAGTGCAGAGAGGCGACGGGCTGAGTATGACAGCCAGTTC 169
QY 117 IleLysLysArgArgLysLeuProSerSerArgArgGlyValSerArgGluGluLeuArg 136
DB 170 ATCAAGAGCGCGCCCTCCCTCCATCCAGCCGGCGGTGTGAGCCGGAGAGATCGAAGC 229
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 230 GAGGTGAGCATCTCGCCGAGATCCGCCACCCCACTATCACTGATGACGATGACGATTC 289
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGlyLeuAspAsp 176
DB 290 GAGAACAGACAGAGATGAGTGTCTGATCTGAGCTGTGCTCGGTGCGAGCTTTTCGAC 349
QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSerArgSer 196
DB 350 TTTCGTGGCGGAGAGAGAGTATGACGAGATGAGCCAGCGAGTTCCTCAACAATC 409
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArg 216
DB 410 CTAGACGGTGTCCACTCTGCTGACTCCAGCGCATCGCACACTTGTGACGAGCCGAG 469
QY 216 gThrSerCysStrTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 470 AACATCATGTTGCTGGAGACAGACAGCCAGCCCGCCGATTAAGCTCATCGACTTGGC 529
QY 236 aSerArgThrGly-SerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSer 256
DB 530 ATCGGCGCACAGATCGAGCTGCGAGAGATTCMAAACAATCTTTGGCA----- 579
QY 256 erSerValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGly 276
DB 579 ----- 579
QY 276 eutThrCysLeuAsnAspProValPheHisSerPro**AspCysLysLeu**ThrThr 296
DB 580 ----- 296
QY 296 rPheLysLysSer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
DB 625 GAGCTTGAGGCTGATGAGCATTTGGCTATCACTACATCACTCTCTGAG----- 677
QY 316 spGlyGlnGlyProGlnThrValProAlaArgGlyProGlyLysArgLysGly----- 333
DB 678 -----CGAGCGTCCCATTTCTGGGCGAGAC--CAAGCAGAGAGACGCTG 720

```

QY 334 -----ThrSerAlaAsnCysLysH 340
DB 721 AGAATCATTCAGACGATGACTATGACTTNTGATGANGAATACTTCAGACAGCAGCGA 780
QY 340 iSTPGLY**GlnMetSerLeuGlyThrLeuAspMetPro 353
DB 781 GCTGGCCAGAGACTTCATCCGACGCTGCTGCTCAAGACCC 821

RESULT 13
BF012103
LOCUS
DEFINITION
BFO12103 696 bp mRNA linear EST 06-OCT-2000
IMAGE:3514032 5' similar to TR:054784 054784 DEATH-ASSOCIATED
KINASE 3 ; , mRNA sequence.

ACCESSION
BF012103
EST.

VERSION
BF012103.1 GI:10712378

KEYWORDS
EST.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

REFERENCE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Other ESTs: U53901.x1

Contact: Robert Strausberg, Ph.D.

Email: cgap@remail.nih.gov

This clone is available royalty-free through LINT ; contact the

IMAGE Consortium (info@image.lint.gov) for further information.

MGI:1396912

Seq primer: -40RP from Gibco

High quality sequence stop: 457.

FEATURES

source

Location/Qualifiers
1. .696
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3514032"
/issue_type="mandible"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="Soares NKMD_mandible"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTTACCAATGTGAAGTGGAGCGCGGCTTATTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 1,746-38 Length: 696
Score: 699.00 Matches: 160
Percent Similarity: 72.53% Conservative: 9
Best Local Similarity: 68.67% Mismatches: 24
Query Match: 38.28% Indels: 40
DB: 10 Gaps: 5

FIGURE9 (1-361) x BF012103 (1-696)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 109 TTTCCTCATCGCGCCAGAGTGCAGAGAGGCGACGGGCTGAGTATGACAGCCAGTTC 168
QY 117 IleLysLysArgArgLysLeuProSerSerArgArgGlyValSerArgGluGluLeuArg 136
DB 169 ATCAAGAGCGCGCTCCATCCAGCCGGCGCGGTGAGCCGGAGAGATCGAAGC 228
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156

```

```

Db      229 GAGGTGAGCATCCGCGGAGATCCGACACCCCAACATCATTAACATGATGATGCTTC 288
Qy      157 GUAenlystrAspValIvalleuIleuGluLeuValSerGlyGlyIleuPheAsp 176
Db      289 GAGAACAGACAGATGAGTGTCTGATCTGGAGCTGGTGTCCGGTGGCGAGCTTTTCGAC 348
Qy      177 PheLeuAaGluLysAsp-His**GlnArgMetArgProArgSerSerSerArgSe 196
Db      349 TTCCTGCGCCGAGAGCATGATGACGAGAGATGAGGCCACGACGCTTCTCAACAATC 408
Qy      196 rTPrThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db      409 CTAGACGATGTCACATCCTGACTCCAGGCGATCCACACTTTCGACCGAAGCCGAG 468
Qy      216 gThrSerCysSerThrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db      469 AACATCATGTTGCTGCAAGACAGCAGCAGCCGCGCCGATTAAGCTCATCGACTTTGGC 528
Qy      236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAaArgLysSerSe 256
Db      529 ATCCGCGACAGAGATCGAGCTGCGACGAGTTCAGAGAACATCTTTGCA----- 577
Qy      256 rSerValArgGlyArgCysGlyHisHisPro-IleGly***IleLeuHisGlyLeuGlyL 276
Db      578 -----CACCGAGTT----- 587
Qy      276 eutThrCysLeuAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
Db      588 -----TGT-----CGCCCGAGATCGGAACACTATGAGCCACTT 621
Qy      296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
Db      622 GGGCTGAGAGCTGACATGAGGAGCAGTGCATCATCACTCAATCTCTCTGAG----- 674
Qy      316 spGlyGlnGlyProGlnThrValProAlaArgly 327
Db      675 -----CGAGGCTGCCCATTCCTCTGG 695

RESULT 14
Bg277312      710 bp      mRNA      linear      EST 21-FEB-2001
LOCUS
DEFINITION
u442b06.y1 Soares NMAX maxillary process Mus musculus cDNA clone
IMAGE:3512794.5, similar to TR:054784 054784 DEATH-ASSOCIATED
KINASE 3 ;, mRNA sequence.
Bg277312
ACCESSION
Bg277312.1 GI:13072489
VERSION
EST.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
1 (bases 1 to 710)
REFERENCE
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Straubeberg, Ph.D.
Other ESTs: ux42b06.x1
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
NCI:339370
Seq primer: -40RP from GILBCO
High quality sequence stop: 477.
Location/Qualifiers
1. 710
FEATURES
Source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3512794"
/tissue_type="maxillary process"
/lab_host="DH10B (phage-resistant)"

```

```

/clone lib="Soares NMAX maxillary process"
/notes="vector: pT733-Pac (pharmacia) with a modified
polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA
was primed with a Not I oligo(CT primer is'
TGTTCACATCTGAGGTGAGAGCGGCCCGCGGCTTTTCTTTTCTTTT 3')
double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT733 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bernaldo."

ORIGIN
Alignment Scores:
Pred. No.:      3,14e-38      Length:      710
Score:          695.50      Matches:      155
Percent Similarity: 75.93%      Conservative: 9
Best Local Similarity: 71.76%      Mismatches: 19
Query Match:     38,098      Indels:      33
DB:              12      Gaps:      4

FIGURE9 (1-361) x Bg277312 (1-710)
Qy      99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db      153 TTTCCTGATGTCGCGAAGTCCAGCAGAGAGGACGGGCGATGATATCGACCAATTC 212
Qy      117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
Db      213 ATCAAGAAAGCGCGCGCTGGCATCCAGCGCGCGGTGAGCGCGGAGAGATCGAACGCC 272
Qy      137 GUAenlystrAspValIvalleuIleuGluLeuValSerGlyGlyIleuPheAsp 156
Db      273 GAGGTGAGATCTCGCGCAGATCCGACCCCAACATCATTAACATGATGACCTGTTTC 332
Qy      157 GUAenlystrAspValIvalleuIleuGluLeuValSerGlyGlyIleuPheAsp 176
Db      333 GAGAACAGACAGATGTGTGCTGATCTCGAGCTGGTGTCCGGTGGGAGCTTTTCGAC 392
Qy      177 PheLeuAaGluLysAsp-His**GlnArgMetArgProArgSerSerSerArgSe 196
Db      393 TTCTCGCGCAGAGAGATGATGAGAGATGAGGCGCAGCCAGCTTCTCAACAATC 452
Qy      196 rTPrThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db      453 CTAGACGGTGTCATCTGACTGACTCCAGCGCATCGACACTTTGACTGAAGCCGAG 512
Qy      216 gThrSerCysSerThrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db      513 ATCATCATGTTTCTGACACAGACGACGACCGCCCTCCGATTAAGCTCATGACATATGAC 572
Qy      236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAaArgLysSerSe 256
Db      573 ATCGGCGACATATGAGGCTGGCGACAGATTCAGAACATCTTTTGA----- 621
Qy      256 rSerValArgGlyArgCysGlyHisHisPro-IleGly**IleLeuHisGlyLeuGlyL 276
Db      622 -----TGT-----CACCGAGTT----- 631
Qy      276 eutThrCysLeuAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
Db      632 -----TGT-----CGTCCCGAGATCGAAGACTATGAGCCACTT 665
Qy      296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHis 310
Db      666 GGCCTTGAAAGCATGATGAGCATTTGGCGGCATCACTCAT 709

RESULT 15
Bg865492      703 bp      mRNA      linear      EST 29-MAY-2001
LOCUS
DEFINITION
602783624P1 NCI_CGAP Mus musculus cDNA clone IMAGE:4910045 5',
ACCESSION
Bg865492
VERSION
Bg865492.1 GI:14216032

```

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 703)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at:
http://image.liml.gov
Plate: LIML0809 row: 0 column: 06
High quality sequence stop: 701.

FEATURES
source 1..703
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4910045"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1; Note: Site 2: Sal; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:
Pred. No.: 7,43e-38 Length: 703
Score: 690.00 Matches: 161
Percent Similarity: 74.78% Conservative: 8
Best Local Similarity: 71.24% Mismatches: 21
Query Match: 37.79% Indels: 37
DB: 12 Gaps: 5

FIGURE9 (1-361) x BG865492 (1-703)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 113 TTGGCATCGTCCGACAGTGCACAGAGGCGACGGGCAATGAGTGCAGCCCAAGTTC 172
QY 117 IleLysLysArgArgLeuProSerSerArgArgLysValSerArgGluGluIleGluArg 136
DB 173 ATCAAGAGCGCGCCCTCCATCCAGCCGCGCGGTGTGAGCCCGGAGAGATGGAACGC 232
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 233 GAGGTGAGCATCTGCGCGAGATCCGCCACCCCAACATCAACACTGACATGACGTGTC 292
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 293 GAGAACAGACAGAGAGTGTGCTGTCTCTGAGCTGTGCTGCTGCTGCTGCTGCTTTCGAC 352
QY 177 PheLeuAlaGluLysAsp-His***GlnArgMetArgProArgSerSerSerSerArgSe 196
DB 353 TTCCTGGCGCGAAGAGACTCATTTGACGAGATGAGGCGACGACGATTCCTCAACAATC 412
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 413 CTAGACGGTGTCCACTTACCTCCACCTCCAGGCGATCGACACTTGAACCTGAAGCCGAG 472
QY 216 gThrSerCysCysrTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 473 AAGATCATGTGTCTGACAGACAGACGACGCCGCCGCCGCAATTAAGCTCATGACTTTGGC 532

```

```

QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 533 ATCGGACACAGATGAGAGCTGTGCGAGCGAGTTCAAGAACATCTTGGCA----- 581
QY 256 rSerValArgGlyArgCysGlyHisHisPro-IleGly**IleLeuHisGlyLeuGlyL 276
DB 582 -----CACCGAGTT----- 591
QY 276 euThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
DB 592 -----TGT-----CGCCCCGAGATCGTGAACATAAGACCACTT 625
QY 296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
DB 626 GGCCTTGAGAGCGTGAACATGTGAGCATTTG-GGTCATCACTTACATCTCTGAGC----- 677
QY 316 sPGLyGlnGlyPro 320
DB 678 --GGAGCGTCCCA 689

```

Search completed: June 30, 2004, 22:20:02
Job time : 3179 secs

XX (UYDU-) UNIV DUKE.
 PA Haystead TA;
 XX WPI; 2002-698619/75.
 PI P-PSDB; A8370856.
 DR
 XX
 PT New myosin phosphatase targeting kinase, useful for identifying compounds
 PT that can treat hypertension or smooth muscle disease.
 XX
 XX Claim 6; Fig 8; 33pp; English.
 XX
 CC The invention relates to myosin phosphatase targeting subunit 1-kinase,
 CC MyPT1 kinase (a subunit of the myosin light chain phosphatase SMP-1M) or
 CC a portion of at least 5 consecutive amino acids. Also included are the
 CC nucleic acid encoding MyPT1 kinase, a MyPT1 kinase expression vector, a
 CC transformed host cell, an anti-MyPT1 kinase antibody, screening a test
 CC compound for anti-hypertensive activity and a kit for detecting the MyPT
 CC kinase comprising a compound that specifically binds to MyPT1 kinase
 CC disclosed within a container means. MyPT1 kinase is useful for identifying
 CC compounds that inhibit hypertensive activity or which treat smooth muscle
 CC disease. The present sequence encodes a MyPT1-kinase from rat aorta
 CC smooth muscle.
 CC
 XX
 XX Sequence 1093 BP; 223 A; 318 C; 318 G; 227 T; 0 U; 7 Other;
 SQ
 Query Match 100.0%; Score 1093; DB 6; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1093; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNTATGNAATATCGGTTTAAATCGGCGGAGCTCGCCNNGGCGAGCTGGAATCCCTCTCA 60
 Db 1 GNTATGNAATATCGGTTTAAATCGGCGGAGCTCGCCNNGGCGAGCTGGAATCCCTCTCA 60
 QY 61 GACCTCTCTCTTCTCGGCTTACAGACGCGGATTAACCTCACTTGAATCTTGGGTCCC 120
 Db 61 GACCTCTCTCTTCTCGGCTTACAGACGCGGATTAACCTCACTTGAATCTTGGGTCCC 120
 QY 121 CGGTGCGCGGCGGAGGTCCTCTCCCTCAAGGCAATCCCAAGTCTGTATGAGGCTCT 180
 Db 121 CGGTGCGCGGCGGAGGTCCTCTCCCTCAAGGCAATCCCAAGTCTGTATGAGGCTCT 180
 QY 181 TTGGGCAATCTGT 240
 Db 181 TTGGGCAATCTGT 240
 QY 241 GTCTGCTTCTCTCTGAGGTCGACGCTTACGCTTCTTCCCTCAAGGCTGCACTTTC 300
 Db 241 GTCTGCTTCTCTCTGAGGTCGACGCTTACGCTTCTTCCCTCAAGGCTGCACTTTC 300
 QY 301 GCCATCGTGCAGAGTGCAGACAGAGGCGACCGGCAATGAGTACGCGGCAATTCATA 360
 Db 301 GCCATCGTGCAGAGTGCAGACAGAGGCGACCGGCAATGAGTACGCGGCAATTCATA 360
 QY 361 AAGAAAGCGGCGCTGCGCTTCCAGCGCGGCTGTAGCGCTGTAGAGATGAGCGCGAG 420
 Db 361 AAGAAAGCGGCGCTGCGCTTCCAGCGCGGCTGTAGCGCTGTAGAGATGAGCGCGAG 420
 QY 421 GTGAGCATCTCGCGGAGATCGGCAACCCCAATCATCATCATCATCATCATCATCATCAT 480
 Db 421 GTGAGCATCTCGCGGAGATCGGCAACCCCAATCATCATCATCATCATCATCATCATCAT 480
 QY 481 AACAGACAGATGT 540
 Db 481 AACAGACAGATGT 540
 QY 541 CTGCTGAGAGATGATCACTGACAGAGATGAGGCGGCAATTCCTTAAGCAATCTCTGG 600
 Db 541 CTGCTGAGAGATGATCACTGACAGAGATGAGGCGGCAATTCCTTAAGCAATCTCTGG 600
 QY 601 ACGGTGTCATCTACTGACTCCAGGCGCATCGGCACTTGAACCTTAAGCGGAGAAC 660
 Db 601 ACGGTGTCATCTACTGACTCCAGGCGCATCGGCACTTGAACCTTAAGCGGAGAAC 660

Db 601 ACGGTGTCATCTACTGACTCCAGGCGCATCGGCACTTGAACCTTAAGCGGAGAAC 660
 QY 661 TCATGTTGCTGAGACATGACAGCAGCCAGCATTAGCTCATCTTGGCATCG 720
 Db 661 TCATGTTGCTGAGACATGACAGCAGCCAGCATTAGCTCATCTTGGCATCG 720
 QY 721 CGACAGATGAGGCGGCTGAGGATTCAGAAATCTTTGGACGCGGAGATTCGTCG 780
 Db 721 CGACAGATGAGGCGGCTGAGGATTCAGAAATCTTTGGACGCGGAGATTCGTCG 780
 QY 781 GTGAGGCGGAGGATGAGGCGGAGCAGCCGATAGGATTTTTCAGGCTTGGCTGAC 840
 Db 781 GTGAGGCGGAGGATGAGGCGGAGCAGCCGATAGGATTTTTCAGGCTTGGCTGAC 840
 QY 841 TGCCTCAACATCTGTCTTCCACAGCCCTGAGATTGTAATGAAACCACTTGGCTT 900
 Db 841 TGCCTCAACATCTGTCTTCCACAGCCCTGAGATTGTAATGAAACCACTTGGCTT 900
 QY 901 GGAAGCTATGATGAGGATGAGGCTGATACACCTACATCTCTGATGAGGCTGAGATGG 960
 Db 901 GGAAGCTATGATGAGGATGAGGCTGATACACCTACATCTCTGATGAGGCTGAGATGG 960
 QY 961 CAGGCGCTCAGACTGTACTGCTAGAGGCGCAGGATCAGGCTGAGCACTTGGCAAC 1020
 Db 961 CAGGCGCTCAGACTGTACTGCTAGAGGCGCAGGATCAGGCTGAGCACTTGGCAAC 1020
 QY 1021 TGCACAACTGCGGCTGAGAGATGTCCTGTGGAACTGTGATGCTGGGCGCCACCA 1080
 Db 1021 TGCACAACTGCGGCTGAGAGATGTCCTGTGGAACTGTGATGCTGGGCGCCACCA 1080
 QY 1081 NGTAGACCATTC 1093
 Db 1081 NGTAGACCATTC 1093
 RESULT 2
 ADBS8463
 ID ADBS8463 standard; DNA; 1514 BP.
 XX
 AC ADBS8463;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3489.
 XX
 KM Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 XX drug screening; toxicity assay; ds.
 OS Unidentified.
 XX
 PN W02003064624-A2.
 XX
 PD 07-AUG-2003.
 XX
 PF 31-JAN-2003; 2003WO-US003194.
 XX
 PR 31-JAN-2002; 2002US-00060087.
 XX
 PR 15-MAR-2002; 2002US-0364045P.
 XX
 PR 30-DEC-2002; 2002US-0436643P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendlick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 XX
 DR WPI; 2003-689530/65.
 XX
 PT Predicting a toxic effect of a compound, useful in identifying toxicity
 PT markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 PT the compound.
 XX
 PS Claim 1; SEQ ID NO 3489; 1156bp; English.


```
OY 408 GATGAGCGGAGGAGGATCTCTGGCGAGATCCGCAACCCCAACATCATCAGCTTGA 467
DB 293 GATGAGCGGAGGAGGATCTCTGGCGAGATCCGCAACCCCAACATCATCAGCTTGA 352
OY 468 CGATGTGTTCGAGAACAAAGACAGATGTGTGTGATCTTGGAGCTGTGTTCGGCGGCA 527
DB 353 CGATGTGTTCGAGAACAAAGACAGATGTGTGTGATCTTGGAGCTGTGTTCGGCGGCA 412
OY 528 ACTTTGACCTTTCTGGCTGAGAGGA-TCACCTACAGAGATGAGGCCAGCACTTCT 596
DB 413 ACTTTGACCTTTCTGGCTGAGAGGA-TCACCTACAGAGATGAGGCCAGCACTTCT 472
OY 587 CAAGAGATCTTGACGCTGTCTCACTACCTGACCTCCAGCGCATCCGCACTTTGACT 646
DB 473 CAAGAGATCTTGACGCTGTCTCACTACCTGACCTCCAGAGCGCATCCGCACTTTGACT 532
OY 647 GAAGCCGAGAACATCATGTGTGTGACACAGATGAGCCAGCCAGCATTTAGCTCAT 706
DB 533 GAAGCCGAGAACATCATGTGTGTGACACAGATGAGCCAGCCAGCATTTAGCTCAT 592
OY 707 CGACTTTGGCATCCGCGCACAGATCGAGGCCGCTAGCGAGTTCAAGAACATCTTTGGCAC 766
DB 593 CGACTTTGGCATCCGCGCACAGATCGAGGCCGCTAGCGAGTTCAAGAACATCTTTGGCAC 652
OY 767 GCCAGAGTTGCTCG 780
DB 653 GCCAGAGTTGCTCG 666
```

RESULT 4

ABN30305 standard; DNA; 65 BP.

```
ID ABN30305
AC AEN30305;
AT 15-JUL-2002 (first entry)
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:3053.
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
OS Rattus norvegicus.
PN WO200210449-A2.
PD 07-FEB-2002.
PF 20-JUL-2001; 2001WO-IB001903.
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
PA (COMP-) COMPUGEN INC.
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
DR
PT New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of a
XX genome, useful for detecting tissue-, pathology-, and developmental-
XX specific genes.
PS Example 1; SEQ ID NO 3053; 47bp; English.
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the (sub-)
XX transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises several
XX oligonucleotides, each capable of hybridizing selectively to a set of
XX messenger RNAs transcribed from a given transcription unit of the genome,
XX which encodes one or more messenger RNA splice variants. The
```

oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterizing the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini-libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue- and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from rats, humans and mice, which are used in the exemplification of the present invention. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from Wipo at http://wipo.int/pub/published_pct_sequences

Query Match 4.0%; Score 44; DB 6; Length 65;

Best Local Similarity 100.0%; Pred. No. 1.2e-11; Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 723 CAAGATCGAGCGCGGTAGCGAGTTCAAGAACATCTTTGGCAC 766
DB 1 CAAGATCGAGCGCGGTAGCGAGTTCAAGAACATCTTTGGCAC 44
```

RESULT 5

AAK34657 standard; DNA; 1429 BP.

```
ID AAK34657
AC AAK34657;
AT 01-JUL-1999 (first entry)
DE Murine ZIP-kinase (serine/threonine kinase) encoding DNA.
KW Zipper Interacting Protein Kinase; ZIP-kinase; serine/threonine kinase;
XX leucine zipper domain; transcription factor Atf4; gene therapy; cancer;
XX Human; murine; ss.
OS Mus musculus.
PN EP911408-A2.
PD 28-APR-1999.
PF 24-SEP-1998; 98EP-00307747.
PR 26-SEP-1997; 97JP-00261589.
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI Akira S, Kawai T;
XX WPI; 1999-246420/21.
DR P-PSDB; AA069322.
PT New Recombinant Zipper Interacting Protein Kinase (ZIP-kinase) protein
XX and DNA, useful as anticancer agents.
PS Claim 6; Page 19-22; 33bp; English.
XX The invention provides human and murine recombinant Zipper Interacting
XX Protein Kinase (ZIP-kinase) proteins. These proteins are serine/threonine
XX kinases which bind the leucine zipper domain of transcription factor
XX Atf4. Host cells containing vectors comprising the ZIP-kinase nucleic
XX acids are used for the recombinant expression of the proteins. ZIP-kinase
XX protein and DNA are useful as gene therapeutic agents against cancer, and
XX as anti-cancer agents. The present sequence represents a DNA encoding a
XX murine ZIP kinase protein.
```

Sequence 1429 BP; 309 A; 425 C; 475 G; 220 T; 0 U; 0 Other;
Query Match 4.0%; Score 44; DB 2; Length 1429;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
DB 196 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 239
RESULT 6
ABL00379/c
ID ABL00379 standard; DNA; 51 BP.
XX
AC ABL00379;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:370.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
OS Homo sapiens.
XX
PN W0200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000MO-US032311.
XX
PR 24-NOV-1999; 99US-0167383P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a pathology,
PT e.g. autoimmune diseases, ascribed to the presence of a sequence
PT polymorphism.
XX
PS Claim 1; Page 360; 674pp; English.
XX
AB100010 to AB101104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56503 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC protein within appropriate physiological samples)
XX
SQ Sequence 51 BP; 11 A; 16 C; 14 G; 10 T; 0 U; 0 Other;
Query Match 2.4%; Score 26; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 911 ATGTGAGCATCGCGGTCAATCCTCA 936

DB 40 ATGTGAGCATCGCGGTCAATCCTCA 15
RESULT 7
AAK93262
ID AAK93262 standard; cDNA; 757 BP.
XX
AC AAK93262;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA clone representative sequence, SEQ ID NO: 1722.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Example 11; SEQ ID NO 1722; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence was used as the representative sequence
CC from a human clone which was used in homology searches to identify the
CC clone. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in CD-ROM format directly from
CC EPO
XX
SQ Sequence 757 BP; 176 A; 219 C; 220 G; 133 T; 0 U; 9 Other;
Query Match 2.4%; Score 26; DB 4; Length 757;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 579 CAGTTCCTCAACAGATCCTGACCG 604
DB 421 CAGTTCCTCAACAGATCCTGACCG 446
RESULT 8
AAK91856
ID AAK91856 standard; cDNA; 757 BP.
XX
AC AAK91856;
XX
DT 06-NOV-2001 (first entry)
XX

DE Human CDNA 5'-end sequence, SEQ ID NO: 316.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 2; SEQ ID NO 316; 1380bp + Sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC method. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO
CC
SQ Sequence 757 BP; 176 A; 219 C; 220 G; 133 T; 0 U; 9 Other;
XX
XX Query Match 2.4%; Score 26; DB 4; Length 757;
XX Best Local Similarity 100.0%; Pred. No. 0.016;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 579 CAGTTCCTCAAGCAGATCCTGACGG 604
DB 421 CAGTTCCTCAAGCAGATCCTGACGG 446
RESULT 9
AAK94258
ID AAK94258 standard; cDNA; 2079 BP.
XX
AC AAK94258;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2874.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
XX
DR F-PSDB; AAM93338.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2874; 1380bp + Sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO
CC
SQ Sequence 2079 BP; 396 A; 626 C; 733 G; 324 T; 0 U; 0 Other;
XX
XX Query Match 2.4%; Score 26; DB 4; Length 2079;
XX Best Local Similarity 100.0%; Pred. No. 0.016;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 579 CAGTTCCTCAAGCAGATCCTGACGG 604
DB 421 CAGTTCCTCAAGCAGATCCTGACGG 446
RESULT 10
ABV72291
ID ABV72291 standard; DNA; 2105 BP.
XX
AC ABV72291;
XX
DT 16-DEC-2002 (first entry)
XX
DE Nucleotide sequence of human DAPK3.
XX
KW Human; IFIT-2; chronic myelogenous leukemia; IAGE-1; BAGE; DDB1; ETS2;
KW PIASy; PIASx-alpha; PIASx-beta; DAPK3; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 94.1458
XX CDS /*tag= a
XX FT /product= "DAPK3"
XX
PN MO200270747-A1.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-JP001901.
XX
PR 01-MAR-2001; 2001JP-00056438.
XX
PA (FUJI) FUJISAWA PHARM CO LTD.
XX
PI Mano H;
XX
DR WPI; 2002-682911/73.

Query Match 2.4%; Score 26; DB 7; Length 2132;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAGCAGATCTCGACGG 604
DB 445 CAGTTCCTCAGCAGATCTCGACGG 470

RESULT 13

AAH16158
ID AAH16158 standard; CDNA: 2224 BP.

AC AAH16158;

DT 26-JUN-2001 (first entry)

DE Human CDNA sequence SEQ ID NO:14925.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2..

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI, 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length CDNA defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length CDNA.

XX Claim 8; SEQ ID NO 14925; 2537bp + Sequence listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length CDNA defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length CDNA. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length CDNA. The primers allow obtaining of the full-length CDNA easily without any specialized methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human CDNA sequences; AA892446 to AA895893 represent human amino acid sequences; and AAH13629 to AAH1632 represent oligonucleotides, all of which are used in the exemplification of the present invention

SQ Sequence 2224 BP; 419 A; 656 C; 806 G; 343 T; 0 U; 0 Other;

Query Match 2.4%; Score 26; DB 4; Length 2224;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CAGTTCCTCAGCAGATCTCGACGG 604
DB 566 CAGTTCCTCAGCAGATCTCGACGG 591

RESULT 14

AAH78068
ID AAH78068 standard; DNA: 2224 BP.

XX AAH78068;

DT 26-NOV-2001 (first entry)

DE Nucleotide sequence of a human protein kinase/protein phosphatase.

XX Human; protein kinase; protein phosphatase; signal transduction;

XX Intracellular signalling pathway; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 215..1579

XX FT /tag= a /product= "protein kinase/protein phosphatase"

XX WQ020109345-A1.

XX 08-FEB-2001.

XX 28-JUL-2000; 2000WO-JP005060.

XX 29-JUL-1999; 99JP-00248036.

XX 18-OCT-1999; 99US-0159590P.

XX 11-JAN-2000; 2000JP-00118776.

XX 17-FEB-2000; 2000US-0183322P.

XX 02-MAY-2000; 2000JP-00183767.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
XX Seno C, Nezu J;

XX WPI, 2001-564736/63.

XX P-PSDB; AA657425.

XX New genes encoding protein kinase and protein phosphatase, useful for identifying mediators which can be used to treat human or animal disorders associated with the expression or function of these enzymes.

XX Claim 1; Page 119-125; 336pp; Japanese.

XX The present sequence encodes a human protein kinase/protein phosphatase. The polypeptides are expected to participate in signal transduction in cells. The kinase phosphatases are connected with intracellular signalling pathways. Antisense oligonucleotides and compounds identified by screening (agonists or antagonists) can be used to treat human or animal disorders associated with the expression or function of the protein. In addition, the polypeptides may be used as target molecules for drug development

XX Sequence 2224 BP; 419 A; 656 C; 806 G; 343 T; 0 U; 0 Other;

Query Match 2.4%; Score 26; DB 5; Length 2224;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

(GENE-) GENE LOGIC INC.
 Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 WPI; 2003-689530/65.
 Predicting a toxic effect of a compound, useful in identifying toxicity markers in liver tissues or cells for drug screening and toxicity assays, comprises preparing gene expression profile of tissue or cells exposed to the compound.
 Claim 1; SEQ ID NO 3720; 1156bp; English.
 The present invention relates to a method for predicting a toxic effect of a compound. The method comprises preparing a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database comprising SEQ ID 1-4925, where differential expression of the gene indicates at least one toxic effect. The method is useful for predicting at least one toxic effect of a compound, predicting hepatotoxicity or the progression of a toxic effect of a compound, identifying an agent that modulates the onset or progression of a toxic response, predicting the cellular pathways that a compound modulates in a cell, and identifying an agent that modulates at least one activity of a protein. The method and compositions of the present invention using a database of genes having liver toxin-induced differential expression, are useful in identifying toxicity markers in liver tissues or cells for drug screening and toxicity assays. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ffp.wipo.int/pub/published_pct_sequences.
 Sequence 2287 BP; 733 A; 446 C; 540 G; 568 T; 0 U; 0 Other;
 Query Match 2.2%; Score 24; DB 9; Length 2287;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 642 GACCTGAAGCCGGAGACATCATG 665
 673 GACCTGAAGCCGGAGACATCATG 696
 RESULT 18
 ADB53383
 ID ADB53383 standard; DNA; 2287 BP.
 XX ADB53383;
 XX
 DT 04-DEC-2803 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3925.
 XX
 KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KW toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN W02003065993-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 04-FEB-2003; 2003WO-US0003482.
 XX
 XX 04-FEB-2002; 2002US-0353171P.
 PR 13-MAR-2002; 2002US-0363534P.
 PR 08-APR-2002; 2002US-0370248P.
 PR 10-APR-2002; 2002US-0371134P.
 PR 10-APR-2002; 2002US-0371135P.
 PR 10-APR-2002; 2002US-0371150P.
 PR 11-APR-2002; 2002US-0371413P.
 PR 19-APR-2002; 2002US-0373601P.
 PR 19-APR-2002; 2002US-0373602P.

22-APR-2002; 2002US-0374139P.
 PR 08-MAY-2002; 2002US-0378370P.
 PR 09-MAY-2002; 2002US-0378652P.
 PR 09-MAY-2002; 2002US-0378653P.
 PR 09-MAY-2002; 2002US-0378655P.
 PR 09-MAY-2002; 2002US-0378656P.
 PR 09-JUL-2002; 2002US-0394230P.
 PR 09-JUL-2002; 2002US-0394233P.
 PR 04-SEP-2002; 2002US-0407688P.
 PR 28-JAN-2003; 2003US-0442900P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
 PI Elashoff M;
 XX
 DR WPI; 2003-731472/69.
 XX
 XX Determining if a compound induces a toxic effect on a tissue or cell, for
 PT identifying hepatotoxic compounds, comprises comparing a gene expression
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
 PT mean values.
 XX
 PS Claim 44; SEQ ID NO 3925; 874pp; English.
 XX
 XX The present invention describes a method for determining whether a
 CC compound induces a toxic effect on a tissue or cell. The method comprises
 CC preparing a gene expression profile of a tissue or cell sample exposed to
 CC the compound, and comparing the gene expression profile to a database
 CC comprising data or information on the Tox mean and non-Tox mean value.
 CC The method is useful for predicting or identifying at least one toxic
 CC effect, particularly hepatotoxicity, of a test or unknown compound. The
 CC genes listed in the specification are useful as diagnostic or toxicity
 CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. These
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.
 CC
 SQ Sequence 2287 BP; 733 A; 446 C; 540 G; 568 T; 0 U; 0 Other;
 Query Match 2.2%; Score 24; DB 9; Length 2287;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 642 GACCTGAAGCCGGAGACATCATG 665
 673 GACCTGAAGCCGGAGACATCATG 696
 RESULT 19
 ACH13101
 ID ACH13101 standard; cDNA; 488 BP.
 XX ACH13101;
 AC ACH13101;
 XX
 DT 13-OCT-2003 (first entry)
 XX
 DE Human adult brain cDNA #313.
 XX
 KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN US2003073623-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 30-JUL-2001; 2001US-00918995.
 XX
 PR 30-JUL-2001; 2001US-00918995.
 XX

PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries; useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 313; 44bp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12769-ACH50831, whose sequence was
CC determined by the technique of SSH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?nocid=20030073623
XX
SQ Sequence 488 BP; 123 A; 125 C; 128 G; 105 T; 0 U; 7 Other;
XX
Query Match 1.9%; Score 21; DB 8; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 919 CATCGCGCTCATCACCCTACAT 939
DB 185 CATCGCGCTCATCACCCTACAT 205
RESULT 20
ABZ77137
ID ABZ77137 standard; cDNA; 1158 BP.
XX
AC ABZ77137;
XX
DT 07-MAY-2003 (first entry)
XX
DE Human protein kinase encoding cDNA SEQ ID NO:23.
XX
KW Human; protein kinase; enzyme; antiasthmatic; antiinflammatory;
KW antidiabetic; antiparkinsonian; antimigraine; cardiant; cystostatic;
KW immunosuppressive; vulnerary; gene therapy; COPD; asthma; migraine;
KW chronic obstructive pulmonary disease; non-insulin dependent diabetes;
KW Parkinson's disease; myocardial infarction; inflammatory bowel disease;
KW autoimmune disorder; allograft rejection; graft versus host disease;
KW cancer; leukaemia; wound granulation; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1158
FT /tag= a
FT /partial
FT /product= "protein kinase"
FT /note= "no start or stop codons given"
XX

PN WO2003000901-A2.
XX
PD 03-JAN-2003.
XX
PF 24-JUN-2002; 2002WO-IB002358.
XX
PR 26-JUN-2001; 2001US-0301098P.
PR 06-NOV-2001; 2001US-0332870P.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Martinez RM, Sigurdsson GT;
XX
DR WPI; 2003-201429/19.
DR F-PSDB; ABP96059.
XX
PT New protein kinase genes and polypeptides, useful for diagnosing diseases
PT associated with a protein kinase, or in gene therapy for treating e.g.
PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
PT or cancers.
XX
PS Claim 1; Page 65; 258bp; English.
XX
CC ABZ77126 to ABZ77165 encode the human protein kinases given in ABP96048
CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cystostatic,
CC immunosuppressive and vulnerary activities, and can be used in gene
CC therapy. A protein kinase therapeutic agent from the present invention,
CC particularly a protein kinase gene agonist or antagonist, can be used for
CC treating a disease or condition associated with a protein kinase in an
CC individual. These diseases include chronic obstructive pulmonary diseases
CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
CC (e.g. leukaemia) or wound granulation
XX
SQ Sequence 1158 BP; 323 A; 290 C; 298 G; 247 T; 0 U; 0 Other;
XX
Query Match 1.9%; Score 21; DB 7; Length 1158;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 919 CATCGCGCTCATCACCCTACAT 939
DB 570 CATCGCGCTCATCACCCTACAT 590
RESULT 21
ABK99973
ID ABK99973 standard; DNA; 1383 BP.
XX
AC ABK99973;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human CADPKL cDNA #1.
XX
KW Human; calcium/calmodulin-dependent protein kinase-like gene; CADPKL; ss;
KW gene; neuropsychiatric disorder; attention deficit disorder; ADD;
KW schizoaffective disorder; bipolar disorder; unipolar affective disorder;
KW schizophrenia; adolescent conduct disorder; pharmacogenetics;
KW fingerprinting; paternity testing; antidepressant; neuroleptic.
XX
OS Homo sapiens.
XX
PN WO200254939-A2.
XX
PD 18-JUL-2002.
XX
PF 07-JAN-2002; 2002WO-US000367.
PF 09-JAN-2001; 2001US-00757300.
PR 23-AUG-2001; 2001US-00935464.
XX

```
XX (MILL-) MILLENNIUM PHARM INC.
PA Meyer JM, Barrington-Martin R, Parker A;
PI WPI; 2002-590643/63.
DR P-PSDB; ABG69792.
XX
XX New variants of calcium/calmodulin-dependent protein kinase-like nucleic
PT acids and polypeptides, useful for diagnosing and treating
PT neuropsychiatric disorders, e.g. schizophrenia, schizoaffective disorder,
PT and bipolar disorder.
XX
XX Claim 1; Page 200-201; 223pp; English.
XX
XX The invention relates to a nucleic acid comprising a polymorphic region
CC of a Calcium/Calmodulin-dependent protein kinase-like gene (CADPKL)
CC allelic variant, and the polypeptide it encodes. CADPKL allelic variants
CC are useful in determining whether a subject has or is at risk of
CC developing a neuropsychiatric disorder, such as schizophrenia, attention
CC deficit disorder (ADD), schizoaffective disorder, bipolar disorder,
CC unipolar affective disorder and adolescent conduct disorder. The
CC polypeptides, polynucleotides, antibodies and modulators of the CADPKL
CC allelic variants are useful for diagnosing or treating these
CC neuropsychiatric disorders. The polypeptides may be used to raise
CC antibodies to a CADPKL polypeptide. The nucleic acids may be used as
CC probes or primers, in pharmacogenomics for designing therapies for the
CC disorders, and in fingerprinting for detection of different individuals
CC with the same species (e.g. paternity testing). This sequence represents
CC human CADPKL cDNA of the invention
XX
SQ Sequence 1383 BP; 355 A; 387 C; 352 G; 289 T; 0 U; 0 Other;
Query Match 1.9%; Score 21; DB 6; Length 1383;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 919 CATCGGCGTCATCACCCTACAT 939
Db 612 CATCGGCGTCATCACCCTACAT 632
RESULT 22
ABT42216
ID ABT42216 standard; DNA; 1388 BP.
XX
XX ABT42216;
AC
XX
XX 26-JUN-2003 (first entry)
DT
XX
XX Toxicity modelling related rat gene SEQ ID No 1918.
DE
XX
XX Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KM database; drug screening; toxicity assay; rat; ds.
XX
XX Rattus norvegicus.
OS
XX
XX WO200295000-A2.
PN
XX
XX 28-NOV-2002.
PD
XX
XX 22-MAY-2002; 2002WO-US016173.
PF
XX
XX 22-MAY-2001; 2001US-0292335P.
PR
XX
XX 13-JUN-2001; 2001US-0297523P.
PR
XX
XX 19-JUN-2001; 2001US-0298925P.
PR
XX
XX 10-JUL-2001; 2001US-0303807P.
PR
XX
XX 10-JUL-2001; 2001US-0303808P.
PR
XX
XX 10-JUL-2001; 2001US-0303810P.
PR
XX
XX 28-AUG-2001; 2001US-0315047P.
PR
XX
XX 27-SEP-2001; 2001US-0324928P.
PR
XX
XX 22-OCT-2001; 2001US-0330462P.
PR
XX
XX 01-NOV-2001; 2001US-0330867P.
PN
```

```
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX
XX Wendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
PI WPI; 2003-148464/14.
DR
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the renal
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
SQ Sequence 1388 BP; 350 A; 369 C; 390 G; 279 T; 0 U; 0 Other;
Query Match 1.9%; Score 21; DB 7; Length 1388;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 642 GACCTGAAGCCGAGACATC 662
Db 524 GACCTGAAGCCGAGACATC 544
RESULT 23
ABK93974
ID ABK93974 standard; cDNA; 1738 BP.
XX
XX ABK93974;
AC
XX
XX 21-OCT-2002 (first entry)
DT
XX
XX Human CADPKL cDNA #2.
DE
XX
XX Human; calcium/calmodulin-dependent protein kinase-like gene; CADPKL; ss;
KM gene; neuropsychiatric disorder; attention deficit disorder; ADD;
KM schizoaffective disorder; bipolar disorder; unipolar affective disorder;
KM schizophrenia; adolescent conduct disorder; pharmacogenomics;
KM fingerprinting; paternity testing; antidepressant; neuroleptic.
OS
XX
XX Homo sapiens.
XX
XX WO200254939-A2.
PN
```


XX (KINE-) KINETEK PHARM INC.
PA (UYER-) UNIV BRITISH COLUMBIA.
XX
XX
PI Yoganathan T, Delaney AD;
DR WPI: 2002-384145/42.
DR P-PSDB; AAE22764.
XX
PT Diagnosing cancer, comprises determining the upregulation of expression
PT of a nucleic acid sequence encoding a protein kinase or upregulation of
PT expression of the protein kinase, in the cancer.
XX
PS Claim 16; Page 62-64; 87pp; English.
XX
CC The invention relates to a method for screening biologically active agent
CC that modulates cancer associated protein kinase function. The invention
CC also relates to a method for diagnosing cancer comprising determining the
CC upregulation of expression of a nucleic acid sequence encoding a protein
CC kinase. The method is useful for diagnosing cancer. A protein kinase is
CC useful for screening biological agents that modulate cancer associated
CC protein kinase function. Downregulating the activity of protein kinase is
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
CC derived tumors and inflammatory samples such as arthritic synovium, for
CC amplified DNA in the cell or increased expression of corresponding mRNA
CC or protein and is also useful to detect differences in expression levels
CC such as molecular weight, amino acid and nucleotide sequences between the
CC two cells. The present sequence is human calmodulin kinase CMK-X1 gene
CC located on chromosome 1q32.1-32.3
XX
SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 U; 0 Other;
Query Match 1.9%; Score 21; DB 6; Length 2447;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 919 CATCGGCGTCATCACCCTACAT 939
Db 681 CATCGGCGTCATCACCCTACAT 701
XX
RESULT 28
AAS31014
ID AAS31014 standard; cDNA; 2689 BP.
XX
AC AAS31014;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide (DITHP) #29.
XX
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
KW respiratory disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200162927-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US006059.
XX
PR 24-FEB-2000; 2000US-0184693P.
PR 24-FEB-2000; 2000US-0184697P.
PR 24-FEB-2000; 2000US-0184698P.
PR 24-FEB-2000; 2000US-0184768P.
PR 24-FEB-2000; 2000US-0184769P.
PR 24-FEB-2000; 2000US-0184770P.
PR 24-FEB-2000; 2000US-0184771P.
PR 24-FEB-2000; 2000US-0184772P.

PR 24-FEB-2000; 2000US-0184773P.
PR 24-FEB-2000; 2000US-0184774P.
PR 24-FEB-2000; 2000US-0184775P.
PR 24-FEB-2000; 2000US-0184776P.
PR 24-FEB-2000; 2000US-0184777P.
PR 24-FEB-2000; 2000US-0184778P.
PR 24-FEB-2000; 2000US-0184813P.
PR 24-FEB-2000; 2000US-0184817P.
PR 24-FEB-2000; 2000US-0184819P.
PR 24-FEB-2000; 2000US-0185213P.
PR 24-FEB-2000; 2000US-0185216P.
PR 12-MAY-2000; 2000US-0203785P.
PR 15-MAY-2000; 2000US-0204226P.
PR 15-MAY-2000; 2000US-0204252P.
PR 16-MAY-2000; 2000US-0204821P.
PR 16-MAY-2000; 2000US-0204908P.
PR 16-MAY-2000; 2000US-0205232P.
PR 17-MAY-2000; 2000US-0204815P.
PR 17-MAY-2000; 2000US-0204863P.
PR 17-MAY-2000; 2000US-0205231P.
PR 17-MAY-2000; 2000US-0205285P.
PR 17-MAY-2000; 2000US-0205286P.
PR 17-MAY-2000; 2000US-0205287P.
PR 17-MAY-2000; 2000US-0205333P.
PR 17-MAY-2000; 2000US-0205324P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Panzer SR, Spito PA, Banville SC, Shah P, Chalup MS, Chang SC,
PI Chen A, D'sa SA, Amsley S, Dahl CR, Dam TC, Daniels SE, Dufour GE,
PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TF,
PI Roseberry AM, Rosen BH, Russo PD, Stockreher TK, Daffo A;
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
XX
DR WPI: 2001-502867/55.
DR P-PSDB; AAU19443.
XX
PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.
XX
PS Claim 1; Page 310-311; 522pp; English.
XX
XX The invention relates to polynucleotides (I) encoding diagnostic and
XX therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and
XX proteins involved in growth and development and receptors. (I) and (II)
XX may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate DITHP expression. For example, (I) and (II)
XX may be used to treat disorders associated with decreased polypeptide
XX expression by rectifying mutations or deletions in a patient's genome,
XX that affect the activity of the DITHPs, by expressing inactive proteins
XX or supplementing the patient's own production of them. (I) and (II) may
XX be used to treat diseases, for example, cell proliferative disorder,
XX Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
XX leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
XX (I) may be used to produce the DITHPs, by inserting the nucleic acids
XX into a host cell and culturing the cell to express the protein. (I) and
XX its complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar nucleic acids in
XX samples, and therefore which patients may be in need of restorative
XX therapy. (II) may also be used as antigens in the production of
XX antibodies against DITHPs and in assays to identify modulators of DITHP
XX expression and activity. The anti-DITHP antibodies and antagonists may
XX also be used to down regulate expression and activity. The anti-DITHP
XX antibodies may also be used as diagnostic agents for detecting the
XX presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay
XX (ELISA)). AAS30966-AAS31196 represent human diagnostic and therapeutic
XX (DITHP) polynucleotides of the invention
XX
SQ Sequence 2689 BP; 642 A; 748 C; 678 G; 621 T; 0 U; 0 Other;
Query Match 1.9%; Score 21; DB 4; Length 2689;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATCGGCGTATCATCCTACAT 939
 DB 848 CATCGGCGTATCATCCTACAT 868

RESULT 29

ABK99972
 ID ABK99972 standard; DNA; 157875 BP.

AC ABK99972;
 XX

DT 21-Oct-2002 (first entry)
 XX

DE Human CADPKL genomic DNA.
 XX

KM Human; calcium/calmodulin-dependent protein kinase-like gene; CADPKL; de;
 KM gene; neuropsychiatric disorder; attention deficit disorder; ADD;
 KM schizophrenia; adolescent conduct disorder; pharmacogenomics;
 KM fingerprinting; paternity testing; antidepressant; neuroleptic.

OS Homo sapiens.
 XX

PN MO200254939-A2.
 XX

PD 18-JUL-2002.
 XX

PF 07-JAN-2002; 2002WC-US000367.
 XX

PR 09-JAN-2001; 2001US-00757300.
 XX

PR 23-AUG-2001; 2001US-00935464.
 XX

PA (MILL-) MILBENNTUM PHARM INC.
 XX

PI Meyer JM, Barrington-Martin R, Parker A;
 XX

DR WPI; 2002-590643/63.
 XX

PT New variants of calcium/calmodulin-dependent protein kinase-like nucleic
 PT acids and polypeptides, useful for diagnosing and treating
 PT neuropsychiatric disorders, e.g. schizophrenia, schizoaffective disorder,
 PT and bipolar disorder.
 XX

PS Claim 1; Page 119-200; 223pp; English.
 XX

CC The invention relates to a nucleic acid comprising a polymorphic region
 CC of a Calcium/Calmodulin-dependent protein kinase-like gene (CADPKL)
 CC allelic variant, and the polypeptide it encodes. CADPKL allelic variants
 CC are useful in determining whether a subject has or is at risk of
 CC developing a neuropsychiatric disorder, such as schizophrenia, attention
 CC deficit disorder (ADD), schizoaffective disorder, bipolar disorder,
 CC unipolar affective disorder and adolescent conduct disorder. The
 CC polypeptides, polynucleotides, antibodies and modulators of the CADPKL
 CC allelic variants are useful for diagnosing or treating these
 CC neuropsychiatric disorders. The polypeptides may be used to raise
 CC antibodies to a CADPKL polypeptide. The nucleic acids may be used as
 CC probes or primers, in pharmacogenomics for designing therapies for the
 CC disorder, and in fingerprinting for detection of different individuals
 CC with the same species (e.g. paternity testing). This sequence represents
 CC human CADPKL genomic DNA of the invention
 XX

Sequence 157875 BP; 44403 A; 35734 C; 34322 G; 43416 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 6; Length 157875;
 Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CATCGGCGTATCATCCTACAT 939

DB 142369 CATCGGCGTATCATCCTACAT 142389

RESULT 30

ABX87988
 ID ABX87988 standard; cDNA; 295 BP.

AC ABX87988;
 XX

DT 24-APR-2003 (first entry)
 XX

DE Corn ear-derived polynucleotide (cpd) #6448.
 XX

KM Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
 KM structural gene; functional gene; regulatory gene;
 KM corn ear-specific profile; gene transcription; gene expression;
 KM hybrid plant; desirable trait expression; plant breeding program;
 KM inheritance; desired characteristic; growth; development;
 KM disease resistance; environmental adaptability; quality; yield;
 KM multigene trait; plant; gene; ss.
 XX

OS Zea mays.
 XX

PN US6476212-B1.
 XX

PD 05-NOV-2002.
 XX

PF 14-MAY-1999; 99US-00313294.
 XX

PR 26-MAY-1998; 98US-0086722P.
 XX

PA (INCY-) INCYTE GENOMICS INC.
 XX

PI Lalgudi RV, Ito LY, Sherman BK;
 XX

DR WPI; 2003-208840/20.
 XX

PT Novel purified corn-ear derived polynucleotide useful as hybridization
 PT probe for detecting polynucleotide in sample, and for identifying,
 PT evaluating, and altering desired characteristics associated with growth,
 PT development.
 XX

PS Example; SEQ ID NO 6448; 390pp; English.
 XX

CC The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
 CC and SATMON023. Some of the cdps uniquely identify structural, functional,
 CC and regulatory genes of corn ear. The polynucleotide sequences are
 CC useful for detecting cdps in a sample, for producing a corn ear-specific
 CC profile of gene transcription, for detecting altered gene expression in
 CC inbred or hybrid plants, and for screening several molecules for specific
 CC binding to the polynucleotide. The cdps are useful to identify, isolate,
 CC or extend identical or related corn-ear nucleic acid sequences from DNA
 CC libraries, and in nucleic acid amplification or hybridization techniques
 CC to follow the expression of desirable traits through plant breeding
 CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with growth
 CC and development, disease resistance, environmental adaptability, quality,
 CC and yield of corn. The cdps are also useful as molecular markers for
 CC studying inheritance and multigene traits in a plant breeding program.
 CC The cdps are useful for producing purified corn-ear polypeptides by
 CC recombinant techniques. They are also useful in diagnostic assays to
 CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cdps) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC segdata.uspto.gov/psipdIDEntry.html
 XX

Sequence 295 BP; 62 A; 83 C; 104 G; 45 T; 0 U; 1 Other;

Query Match 1.8%; Score 20; DB 7; Length 295;
 Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 TCCGCCACCCCAATCATC 459

Db 157 TCCGCCACCCCAATCATC 176

RESULT 31

ABAS8634
ID ABAS8634 standard; DNA; 447 BP.

XX
AC ABAS8634;

XX
DT 01-FEB-2002 (first entry)

XX
DE Human foetal liver single exon nucleic acid probe #6939.

XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX
OS Homo sapiens.

XX
PN WO200157277-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000669.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

XX
PR 27-SEP-2000; 2000US-0236359P.

XX
PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-483447/52.

XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX
PT gene expression in human fetal liver.

XX
PS Claim 1; SEQ ID NO 6939; 639bp + Sequence Listing; English.

XX
SQ Sequence 447 BP; 95 A; 140 C; 74 G; 138 T; 0 U; 0 Other;

XX
Query Match 1.8%; Score 20; DB 4; Length 447;

XX
Best Local Similarity 100.0%; Pred. No. 18;

XX
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 TGATCTTGAGCTGCTGCC 519

Db 284 TGATCTTGAGCTGCTGCC 303

RESULT 32

AA18308
ID AA18308 standard; DNA; 447 BP.

XX
AC AA18308;

XX
DT 17-OCT-2001 (first entry)

XX
DE Probe #6994 used to measure gene expression in human placenta sample.

KW Probe; microarray; human; placenta; antenatal diagnosis;

XX
KW genetic disorder; ss.

XX
OS Homo sapiens.

XX
PN WO200157272-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000663.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

XX
PR 27-SEP-2000; 2000US-0236359P.

XX
PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-488897/53.

XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX
PT gene expression in human placenta.

XX
PS Claim 25; SEQ ID NO 6994; 654bp; English.

XX
CC The present invention relates to single exon nucleic acid probes (SENP).

XX
CC The present sequence is one such probe. The probes are useful for

XX
CC producing a microarray for predicting, measuring and displaying gene

XX
CC expression in samples derived from human placenta. The probes are useful

XX
CC for antenatal diagnosis of human genetic disorders

XX
SQ Sequence 447 BP; 95 A; 140 C; 74 G; 138 T; 0 U; 0 Other;

XX
Query Match 1.8%; Score 20; DB 4; Length 447;

XX
Best Local Similarity 100.0%; Pred. No. 18;

XX
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 TGATCTTGAGCTGCTGCC 519

Db 284 TGATCTTGAGCTGCTGCC 303

RESULT 33

AAK32477
ID AAK32477 standard; DNA; 447 BP.

XX
AC AAK32477;

XX
DT 06-NOV-2001 (first entry)

XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 7034.

XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX
OS Homo sapiens.

XX
PN WO200157276-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000668.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 7034; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
XX Sequence 447 BP; 95 A; 140 C; 74 G; 138 T; 0 U; 0 Other;
SQ
Query Match 1.8%; Score 20; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 500 TGATCTTGAGCTGCTGCC 519
Db 284 TGATCTTGAGCTGCTGCC 303
RESULT 34
AAK06764
ID AAK06764 standard; DNA; 447 BP.
XX
XX AAK06764;
AC
XX
XX 05-NOV-2001 (first entry)
DT
XX Human brain expressed single exon probe SEQ ID NO: 6755.
DE
XX Human brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KM ss.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000667.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
XX Example 4; SEQ ID NO 6755; 650bp + Sequence Listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
XX Sequence 447 BP; 95 A; 140 C; 74 G; 138 T; 0 U; 0 Other;
SQ
Query Match 1.8%; Score 20; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 500 TGATCTTGAGCTGCTGCC 519
Db 284 TGATCTTGAGCTGCTGCC 303
RESULT 35
ABS32185
ID ABS32185 standard; DNA; 447 BP.
XX
XX ABS32185;
AC
XX
XX 25-FEB-2003 (first entry)
DT
XX Human liver single exon probe, SEQ ID NO 7175.
DE
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KM coronary heart disease; ss.
XX
XX Homo sapiens.
OS
XX WO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000664.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 7175; 658bp; English.
PS
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 1109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 447 BP; 95 A; 140 C; 74 G; 138 T; 0 U; 0 Other;

Query Match 1.8%; Score 20; DB 4; Length 447;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 TGATCTTGAGCTGTGTC 519
 |||||
 Db 284 TGATCTTGAGCTGTGTC 303

RESULT 36
 ABS07263
 ID ABS07263 standard; DNA; 447 BP.

XX ABS07263;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID No 7254.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; IHD;

KW Chronic obstructive pulmonary disease; interstitial lung disease;

KM familial idiopathic pulmonary fibrosis; neurofibromatosis;

KM Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;

KM pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;

KM primary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KM primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

PD 30-JAN-2001; 2001WO-US000665.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples.

XX Claim 1; SEQ ID NO 7254; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of probes

CC; the novel set of probes which hybridise at high stringency to a nucleic

CC acid expressed in the human lung; measuring gene expression in a sample

CC derived from human lung; comprising (a) contacting the array with a

CC collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the

CC array; identifying exons in a eukaryotic genome, comprising (a)

CC algorithmically predicting at least one exon from genomic sequences of

CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Heremansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemorrhoidosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 447 BP; 95 A; 140 C; 74 G; 138 T; 0 U; 0 Other;
 SQ
 Query Match 1.8%; Score 20; DB 6; Length 447;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 TGATCTTGAGCTGTGTC 519
 |||||
 Db 284 TGATCTTGAGCTGTGTC 303

RESULT 37
 ACHS0827
 ID ACHS0827 standard; cDNA; 493 BP.

XX ACHS0827;

DT 13-OCT-2003 (first entry)

XX Human mammary gland cDNA #232.

DE

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

PD 30-JUL-2001; 2001US-00918995.

PR 30-JUL-2001; 2001US-00918995.

XX (DBMA/) DRMANAC R T.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON W C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson WC, Jones LW;

XX WPI; 2003-615964/58.

DR New polynucleotide sequences obtained from various cDNA libraries, useful

PT as hybridisation probes, as oligomers for PCR, for chromosome and gene

XX mapping, in the recombinant production of protein, or in generating

```

PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 38039; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030073623
XX
SQ Sequence 493 BP; 126 A; 118 C; 125 G; 121 T; 0 U; 3 Other;

Query Match          1.8%; Score 20; DB 8; Length 493;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 923 GGGCTCATCACCCTCATCCT 942
Db 471 GGGCTCATCACCCTCATCCT 490

RESULT 38
AAT17177
ID AAT17177 standard; DNA; 748 BP.
XX
AC AAT17177;
XX
DT 24-OCT-1996 (first entry)
XX
DE Rhodococcus rhodochrous KM-02 kanamycin resistance gene.
XX
KM KM-02 strain; kanamycin resistance gene; marker; construction;
XX self cloning system; ds.
XX
OS Rhodococcus rhodochrous.
XX
FH Key Location/Qualifiers
FT CDS 165..680
FT /*tag= a
XX
XX EP704530-A2.
XX
XX 03-APR-1996.
XX
XX 04-AUG-1995; 95BP-00112298.
XX
XX 04-AUG-1994; 94JP-00201582.
XX
XX (NITTO ) NITTO CHEM IND CO LTD.
XX
XX Yu F, Kato M;
XX
XX WPI, 1996-173034/18.
XX
XX P-PSDB; AAR33155.
XX
XX Rhodococcus sp. Kanamycin resistance gene - useful as a marker for
XX construction of self-cloning systems.
XX
XX Claim 2; Page 10; 15pp; English.
XX
XX The present sequence is the R. rhodochrous, strain KM-02, kanamycin

```

```

CC resistance gene, useful as a marker in the construction of a self cloning
CC system, pref. for the self cloning of R. rhodochrous. Only the CDS of the
CC gene is claimed, however the entire gene sequence is referred to in
CC example 1 of the specification
XX

SQ Sequence 748 BP; 138 A; 240 C; 225 G; 145 T; 0 U; 0 Other;

Query Match          1.8%; Score 20; DB 2; Length 748;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 658 ACATCATGTTGCTGACCAAG 677
Db 708 ACATCATGTTGCTGACCAAG 727

RESULT 39
ABA09608
ID ABA09608 standard; DNA; 1120 BP.
XX
AC ABA09608;
XX
DT 15-JAN-2002 (first entry)
XX
DE Human bone marrow expressed oligonucleotide SEQ ID NO: 34.
XX
XX
XX Human bone marrow; cytostatic; antineumatic; antiarthritic; vulnary;
XX antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
XX antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
XX anticancer; fungicide; antidiabetic; antistatic; antiallergic;
XX immunostimulant; analgesic; cerebroprotective; antianemic; infection;
XX nervous system disorder; autoimmune disorder; inflammation; allergy; ds.
XX
OS Homo sapiens.
XX
XX WO200174836-A1.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US010472.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX 30-NOV-2000; 2000US-0250583P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dermanac RT, Ford JE, Boyle BJ;
XX
XX WPI; 2001-626375/72.
XX
XX P-PSDB; ABB12364.
XX
XX New bone marrow-expressed nucleic acids and polypeptides, useful for
XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
XX and increasing hematopoiesis, stem cell survival and bone growth and
XX remodeling.
XX
XX Claim 1; Page 182; 380pp; English.
XX
XX The present invention relates to bone marrow expressed polynucleotides
XX and proteins. These sequences can be used in the treatment of
XX inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
XX and peripheral nervous system diseases and neuropathies, such as
XX Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
XX disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
XX cell disorders, platelet disorders, stem cell disorders, bone
XX degenerative disorders, autoimmune disorders, for example multiple
XX sclerosis, diabetes and arthritis, viral and bacterial infections,
XX allergies and blood coagulation disorders. The present sequence is a DNA
XX
XX Sequence 1120 BP; 271 A; 304 C; 315 G; 230 T; 0 U; 0 Other;

```

Query Match 1.8%; Score 20; DB 5; Length 1120;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 923 GGCGTCATCAGCTACATCCT 942
DB 58 GGCGTCATCAGCTACATCCT 77

RESULT 40

ADA69943
ID ADA69943 standard; DNA; 1130 BP.

XX
AC ADA69943;

DT 20-NOV-2003 (first entry)

XX
DE Rice gene, SEQ ID 3266.

XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
gene; ds.

XX
OS Oryza sativa.

XX
PN WO200300898-A1.

XX
PD 03-JAN-2003.

XX
PF 22-JUN-2001; 2001WO-IB001105.

XX
PR 22-JUN-2001; 2001WO-IB001105.

XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katsagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

XX
DR WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX
PS Claim 6; SEQ ID NO 3266; 899pp; English.

XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX
SQ Sequence 1130 BP; 272 A; 305 C; 303 G; 250 T; 0 U; 0 Other;

Query Match 1.8%; Score 20; DB 7; Length 1130;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 TCCGCCACCCCAACATCATC 459
DB 359 TCCGCCACCCCAACATCATC 378

RESULT 41

AAS73461
ID AAS73461 standard; CDNA; 1253 BP.

XX
AC AAS73461;

XX
DT 13-FEB-2002 (first entry)

XX
DE DNA encoding novel human diagnostic protein #265.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX
OS Homo sapiens.

XX
PN WO200175067-A2.

XX
PD 11-OCT-2001.

XX
PF 30-MAR-2001; 2001WO-US008631.

XX
PR 31-MAR-2000; 2000US-00540217.

XX
PR 23-AUG-2000; 2000US-00649167.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Drmanac RT, Liu C, Tang YT;

XX
DR WPI; 2001-639362/73.

XX
DR P-PSDB; ABS09274.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX
PS Claim 1; SEQ ID NO 9265; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

XX
SQ Sequence 1253 BP; 331 A; 303 C; 366 G; 253 T; 0 U; 0 Other;

Query Match 1.8%; Score 20; DB 5; Length 1253;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 AGCCGCGCGGTGTGAGCCG 401
DB 324 AGCCGCGCGGTGTGAGCCG 343

RESULT 42

AAC46591
ID AAC46591 standard; DNA; 1435 BP.

XX
AC AAC46591;

XX
DT 18-OCT-2000 (first entry)

XX Zea mays DNA fragment SEQ ID NO: 50700.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic; pathway;
KM promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EPI033405-A2.
XX
FD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01233180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 09-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 21-APR-1999; 99US-0129845P.
PR 23-APR-1999; 99US-0130077P.
PR 28-APR-1999; 99US-0130449P.
PR 30-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137828P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138864P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144633P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145122P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148177P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.

PR	27-AUG-1999	99US-0151080P
PR	30-AUG-1999	99US-0151303P
PR	31-AUG-1999	99US-0151438P
PR	01-SEP-1999	99US-0151929P
PR	07-SEP-1999	99US-0151836P
PR	10-SEP-1999	99US-0153070P
PR	13-SEP-1999	99US-0153758P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154779P
PR	22-SEP-1999	99US-0155139P
PR	22-SEP-1999	99US-0155488P
PR	24-SEP-1999	99US-0155659P
PR	28-SEP-1999	99US-0156458P
PR	29-SEP-1999	99US-0156596P
PR	04-OCT-1999	99US-0157117P
PR	05-OCT-1999	99US-0157753P
PR	06-OCT-1999	99US-0157865P
PR	07-OCT-1999	99US-0158028P
PR	08-OCT-1999	99US-015832P
PR	12-OCT-1999	99US-0158369P
PR	13-OCT-1999	99US-0159293P
PR	13-OCT-1999	99US-0159294P
PR	14-OCT-1999	99US-0159295P
PR	14-OCT-1999	99US-0159329P
PR	14-OCT-1999	99US-0159330P
PR	14-OCT-1999	99US-0159331P
PR	14-OCT-1999	99US-0159637P
PR	14-OCT-1999	99US-0159638P
PR	18-OCT-1999	99US-0159584P
PR	21-OCT-1999	99US-0160741P
PR	21-OCT-1999	99US-0160767P
PR	21-OCT-1999	99US-0160768P
PR	21-OCT-1999	99US-0160770P
PR	21-OCT-1999	99US-0160814P
PR	21-OCT-1999	99US-0160815P
PR	22-OCT-1999	99US-0160980P
PR	22-OCT-1999	99US-0160981P
PR	22-OCT-1999	99US-0160989P
PR	25-OCT-1999	99US-0161404P
PR	25-OCT-1999	99US-0161405P
PR	25-OCT-1999	99US-0161406P
PR	26-OCT-1999	99US-0161359P
PR	26-OCT-1999	99US-0161360P
PR	26-OCT-1999	99US-0161361P
PR	28-OCT-1999	99US-0161920P
PR	28-OCT-1999	99US-0161992P
PR	28-OCT-1999	99US-0161993P
PR	29-OCT-1999	99US-0162142P


```

FT XX /tag= b
XX PN WO9966030-A1.
XX PD 23-DEC-1999.
XX PF 15-JUN-1999; 99WO-US013411.
XX PR 15-JUN-1998; 98US-0089234P.
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PA (MCIN/) MCINNIS P A.
XX PI Kimchi A;
XX DR WPI: 2000-147148/13.
XX DR P-PSDB; AAY44674.
XX PT Calmodulin-dependent serine/threonine kinase capable of inducing
XX PT apoptosis used in the treatment of e.g. cancer.
XX PS Claim 3; Fig 1; 67pp; English.
XX CC The present sequence encodes DAP (death-associated protein) - kinase-
XX CC related protein 1 (DRP-1), which is a calmodulin-dependent
XX CC serine/threonine kinase. DRP-1 is a cytoplasmic protein capable of
XX CC inducing apoptosis by dimerisation. It shows significant homology to DAP
XX CC kinase. It has cytostatic, antiproliferative and immunosuppressive activity
XX CC and can be used for inhibiting growth/metastasis of tumours and promoting
XX CC death of tumour cells. It can also be used in the treatment of cancer,
XX CC psoriasis and autoimmune diseases. Fragments of DRP-1 DNA are useful as
XX CC probes for screening individuals with a predisposition to cancer
XX SQ Sequence 1742 BP; 442 A; 437 C; 504 G; 359 T; 0 U; 0 Other;

Query Match 1.8%; Score 20; DB 3; Length 1742;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 AGCGGCGCGGTGAGCCG 401
Db 215 AGCGGCGCGGTGAGCCG 234

RESULT 45
AAV49655
ID AAV49655 standard; DNA; 28720 BP.
XX AC AAV49655;
XX DT 23-OCT-1998 (first entry)
XX DE Human SC3 DNA.
XX SC Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
XX KM nervous system; medullo-blastoma; glioma; breast; detection; SC3;
XX KM autoantibody; se.
XX OS Homo sapiens.
XX PN WO9830687-A2.
XX PD 16-JUL-1998.
XX PF 09-JAN-1998; 98WO-DE000096.
XX PR 09-JAN-1997; 97DE-01000519.
XX PR 18-JUL-1997; 97DE-01030997.
XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PI Mollenhauer J, Poustka A;
XX

```

```

DR WPI: 1998-399136/34.
XX PT Proteins containing scavenger receptor, cysteine rich domain - useful for
XX PT diagnosis and treatment of tumours.
XX PS Disclosure; Fig 5c; 54pp; German.
XX CC This nucleotide sequence contains a fragment of a gene which is capable
XX CC of encoding a protein which contains a SRCR (scavenger receptor, cysteine
XX CC rich) domain. The gene and encoded protein can be used to diagnose or
XX CC treat tumours, particularly of the nervous system (medullo-blastoma or
XX CC glioma) or breast. The DNA sequence and probes derived from it, are used
XX CC to identify genes that express SRCR-domain containing proteins, to
XX CC determine the form in which these proteins exist and to assess the
XX CC significance of individual forms on cellular properties. The protein can
XX CC be used to detect the presence of autoantibodies and antibodies which
XX CC regulate its expression
XX SQ Sequence 28720 BP; 6956 A; 6728 C; 6982 G; 8054 T; 0 U; 0 Other;

Query Match 1.8%; Score 20; DB 2; Length 28720;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 952 TGAGATGGGCGAGGCGCTCA 971
Db 22565 TGAGATGGGCGAGGCGCTCA 22584

Search completed: June 30, 2004, 19:02:41
Job time : 512 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 01:24:35 ; Search time 6123 Seconds
(without alignments) 2555.420 Million cell updates/sec

2555.420 Million cell updates/sec

Title: FIGURE9
Sequence: 1 XM1G1IGRSSPXQLDLS.....GXEMSLGTLMDPQHQRXTX 361

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgnt 1/USPFO.spool/GBRREYESUS641/runat 29062004 143151 24560/app_query.fasta_1.519
-DB=genembl -QMT=fastap -SUFPRfx=rgc -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -WATRX-Biosum62 -TRANS=humat40.cdt -LIST=45
-DOCALLIGN=200 -THR_SCORE=spc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USFR=GBRREYESUS641 @CGN 1 1 3508 @runat 29062004 143151 24560 -NCPU=6 -ICPU=3
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_in:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	1533	84.0	24009	2 AC120292	AC120292 Rattus no
2	1513.5	82.9	218817	2 AC114704	AC114704 Rattus no
3	1513.5	82.9	246796	2 AC094643	AC094643 Rattus no
4	930	50.9	247196	2 AC073822	AC073822 Mus muscu
5	858.5	47.0	1560	10 BC062076	BC062076 Rattus no
6	856.5	46.9	1514	10 RMO6971	AM006971 Rattus no
7	722	39.5	1410	10 AB007143	AB007143 Mus muscu
8	722	39.5	1429	6 AR076190	AR076190 Sequence
9	722	39.5	1429	6 AR124103	AR124103 Sequence
10	722	39.5	1429	6 E23385	E23385 DNA encodin
11	632	34.6	1365	12 AY335738	AY335738 Synthetic
12	632	34.6	2055	6 AB022341	AB022341 Homo sapi
13	632	34.6	2079	6 BD127287	BD127287 Primer fo
14	632	34.6	2079	6 AK074799	AK074799 Homo sapi
15	632	34.6	2105	6 BD176607	BD176607 Method of
16	632	34.6	2105	6 AB007144	AB007144 Homo sapi
17	632	34.6	2132	6 AR076189	AR076189 Sequence
18	632	34.6	2132	6 AR124102	AR124102 Sequence
19	632	34.6	2132	6 E23384	E23384 DNA encodin
20	631.5	34.6	2224	6 AX880020	AX880020 Sequence
21	631.5	34.6	2224	6 BD012208	BD012208 Novel gen
22	631.5	34.6	2224	6 BD158150	BD158150 Primer fo
23	631.5	34.6	2224	9 AK027590	AK027590 Homo sapi
24	607.5	33.3	2118	9 AK097643	AK097643 Homo sapi
25	581.5	31.8	757	6 BD124885	BD124885 Primer fo
26	581.5	31.8	757	6 BD126291	BD126291 Primer fo
27	565.5	31.0	210617	2 AC034201	AC034201 Homo sapi
28	548	30.0	134308	9 AC011488	AC011488 Homo sapi
29	500.5	27.4	138542	2 AC113577	AC113577 Tetradon
30	469.5	25.7	1739	9 AB018001	AB018001 Homo sapi
31	469.5	25.7	1742	9 AF052941	AF052941 Homo sapi
32	453	24.8	1883	10 BC022165	BC022165 Mus muscu
33	445.5	24.4	4910	10 AY245541	AY245541 Mus muscu
34	445.5	24.4	5003	10 AY245540	AY245540 Mus muscu
35	445.5	24.4	5056	10 BC057317	BC057317 Mus muscu
36	445.5	24.4	5056	10 BC060161	BC060161 Mus muscu
37	445.5	24.4	5345	10 NM0486	X97048 M. musculus
38	443	24.3	197275	2 BX539306	BX539306 Dantio rer
39	443	24.3	234956	2 BX539340	BX539340 Mus muscu
40	437	23.9	1757	10 AB018002	AB018002 Homo sapi
41	432.5	23.7	1104	9 BT006935	BT006935 Homo sapi
42	432.5	23.7	1104	12 BT007780	BT007780 Synthetic
43	432.5	23.7	5910	9 HSDAPK	X76104 H. sapiens D
44	428.5	23.5	4272	6 BD193362	BD193362 Composite
45	428.5	23.5	4935	6 AR080622	AR080622 Sequence

RESULT 1

ALIGNMENTS

AC120292	240009 bp	DNA	linear	HTG 15-NOV-2002
LOCUS				
DEFINITION	Rattus norvegicus clone CH230-23B15, WORKING DRAFT SEQUENCE, 2			
ACCESSION	AC120292			
VERSION	AC120292.4			
KEYWORDS	HTG; HTGS; PHASE1; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE	Rattus norvegicus			
ORGANISM	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	1 (bases 1 to 240009)			
AUTHORS	Munzy,D,Marie., Metzker,M,lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Albrocks,S., Amn,A., Angliano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandranaika,D., Barber,M., Barnstead,M., Benahmed,F., Biwalto,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Cesaar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Ande,C., Dederich,D., Delgado,O., Denison,S., Detamo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durkin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,T., Foster,P., Fraser,C.M., Gabisi,A., Gantla,R., Garcia,A., Garner,T., Garza,M., Garagaris,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulvyk,S., Hume,G., Idsbit,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Kapaty,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,Y., Liu,Y., London,P., Longacre,S., Lopez,J., Loutensheha,L., Loulsegad,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martini,K., Martin,R., Martinez,E., Mathiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Mnjaja,E., Montemayor,E., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L., Narkervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwekeleleh,O., Okunnu,G., Olarpunagoon,A., Pal,S., Parke,K., Pasterkamp,S., Paul,H., Perez,A., Perez,L., Pfamkoch,C., Plodder,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Kelly,B., Kelly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tringey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleciwyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.			
TITLE	Direct Submission			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 240009)			
AUTHORS	Worley,K.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA			
REFERENCE	3 (bases 1 to 240009)			
AUTHORS	Rat Genome Sequencing Consortium.			
TITLE	Direct Submission			

```

JOURNAL
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23609710.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GWGC
Center clone name: CH230-23B15
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 207207 bases at least Q40
Consensus quality: 208683 bases at least Q30
Consensus quality: 209634 bases at least Q20
Estimated insert size: 213115; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html) .
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 223747: contig of 223747 bp in length
* 223748 223847: gap of unknown length
* 223848 240009: contig of 16162 bp in length.
Location/Qualifiers
1..240008
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-23B15"
complement(223684..223551)
/note="clone_boundary"
clone_end:T7
site:
end_sequence:B2106119"
complement(222705..223551)
/note="clone_boundary"
clone_end:T7
site:
end_sequence:B2106119"

FEATURES
Source
misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Alignment Scores:
Pred. No.: 3,53e-59 Length: 240009
Score: 1533.00 Matches: 327
Percent Similarity: 66.87% Conservative: 2
Percent Local Similarity: 66.46% Mismatches: 13
Query Match: 83.95% Indels: 150
DB: 2 Gaps: 4

FIGURE9 (1-361) x AC120292 (1-240009)

```

```

QY 12 Pro***GlyGlnLeuAspSerLeuSerAspLeuLeuSerAProGlnHISGlyLeu 31
Db 213636 CCGTCAGAGGAGCGGAGCTCCCTCTCAGACCTCCTTTCTCTCCCTCAGCAGGAGAT 213695
QY 32 AsnLeuThr***LeuPheLeuGlyProArgCysArgLaserValLeuSerLeuLysAla 51
Db 213696 AACCTCACTTGAGCTGCTTGGGGTCCCGGTCGCGGCGCAGGCTCCTCCCTCAAGGCA 213755
QY 52 IleProLysCysLeuSer***GlySerLeuGlySerSerValValGlyAsnLeuGly 71
Db 213756 ATCCCAAGAGTGTCTGATGAGGCTCTTGGGCAAGTCTGTTGTGGGAAACCTGGGA 213815
QY 72 ThrAspAlaGlnAspLeuGlyTyrArgValLeuProSerSerGlySerAlaAlaLeuSer 91
Db 213816 ACAGATGCACAGAGCGTGGGGTACAGAGTCTCCTGCTCTGAGTGTGAGCGCTTAGC 213875
QY 92 CysSerPheProHisSerGly---PheAlaIleValArgLysCys-----LysGlyThr 108
Db 213876 TGTTCCTTCCCGCCACAGCGGCGCATGTCGCGCAGAGTGCAGAGGAGGAGGACACC 213935
QY 109 GlyMetGlnTyrAlaAlaLysPheIleValSerArgLysLeuProSerSerArgArgGly 128
Db 213936 GCGATGAGATACGCGCGCAAGTTCATAAAGAGCGGCGCTGCCGCTCCAGCGCGCGGT 213995
QY 129 ValSerArgGlnGlnIleGlnArgLysValSerIleLeuArgLysIleArgHisProAsn 148
Db 213996 GTGAGCGGTGAGGAGATCGAGCGGAGGTGAGCATCTGCGCGAGATCCGCCACCCCAAC 214055
QY 149 IleIleThrLeuHisAspValPheGlnAsnLysThrAspValValLeuIleLeuGlnLeu 168
Db 214056 ATCATTCACGCTGCACGATGTGTTCTGAGAACAGACATATGTGGTGTGAGAGCTG 214115
QY 169 ValSerGlyGlyGlnLeuPheAspPheLeuAlaGlyLysAsp-His***GlnArgMetArg 188
Db 214116 GTGTCCGGCGCGGAGACTTTTGACTTCTGCTGAGAGAGAGTACACGACAGAGAGT 214175
QY 188 GProArgSerSerSerSerSerArgSerTTPThrValSerThrThCysThrProSerAlaSe 208
Db 214176 GCCACGCGATGCTCTCAAGACATCTGACGAGTGTCCACTCACTGACCTCAAGCGCATC 214235
QY 208 PArgThrLeuThr----- 212
Db 214236 GCGGACTTTTGAAGTGAAGTGAAGTGTGTGTGATGAGGCGGTGGAGGAGGAGGAGGATA 214295
QY 213 ----- 222
Db 214296 ACCAAGGAGACACACACACACCGCGCTCTCGCAGCGGAGAACATCATGTTGCTGAGC 214355
QY 222 rSerMetGlnProIleHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerArg 242
Db 214356 AAGCATGCAGCCACCGACCGATTAAGCTCATGACTTGGCAGCGCGCAGAGATCGAG 214415
QY 242 GProValAlaSerSerArgThrSerLeuAlaArgGlnSerSerSerValArgLysArgCys 262
Db 214416 GCGCGTACGCGAGTTCAGAACATCTTTGGACCGCAGAGTTCGTCGCTGAGGGGCGAGGTG 214475
QY 262 sGlyHisHisProIleLeu***IleLeuHisGlyLeuGlyLeuThrCysLeuAsnArg 282
Db 214476 TGGGACACACCGCATAGGATGATTCGACCGCGCTTGGCTGACCTCAACATCC 214535
QY 282 cValPheHisSerPro***AspCysLysLeu***ThrThrTyrLeuGlySer***TyrVal 302
Db 214536 TGTCTTCCACAGCCCGCTGAGATTGTAAACTATGAACCACTTGGCTGGAAAGCTGATATGT 214595
QY 302 1----- 302
Db 214596 GGTGAGGTGGGTTGGAGAGCGCTGCTCCATCTCTCCGTCGTCCTCCGTCCT 214655
QY 302 ----- 302
Db 214656 CCGCTTAAGTGTCCCAACCTGTGGGCTTATTCAGCTAGTAGAGGTGTGGGCGCTAGACA 214715
QY 302 ----- 302

```

```

Db 214716 GCAGCATGGAAGGCGCTGAGACCCCTGCTCAGACCCCAACTATCAGACTATAGGGCTC 214775
QY 302 ----- 302
Db 214776 TGCTCAGTCACAGCGGCGCTTGTGTCGTGCGATAGAGGTCTGAGGTTCTGCTGTGCG 214835
QY 302 ----- 302
Db 214836 TGCCAGAGTGTCTCAGGCAACCTGTGTTGACACTGACACACACCTTAGAGTGGGCGCTA 214895
QY 303 ----- 303
Db 214896 TGTCCTGGGGGTGAGTGTGTCGCCAAGCGCATCTCTACACAGACTTTCTGCTCCCTTGACAGA 214955
QY 303 UHisArgArgHisHisLeuHisProValSerAla***AspGlyGlnGlyProGlnThrVal 323
Db 214956 GCATGGCGCTCATCCCTACCTACATCCCTGTGATGCTGAGATGGGAGGCGGCGCTCAGACTGT 215015
QY 323 lProAlaArgGlyProGlyIleArgAlaGlyThrSerAlaAsnCysLysHisTyrGly** 343
Db 215016 ACCTGCTAGAGGCCCAAGGATCAGGCGCTGAGCACTCTCCAAACTGCAAACTGGCGCTG 215075
QY 343 *GlnMetSerLeu-GlyThrLeuAspMetPro 353
Db 215076 AGAGATCTCCCTGGGGAACACTGCGCTATGCTT 215107

RESULT 2
AC114704/c 218817 bp DNA linear HTG 20-NOV-2002
LOCUS AC114704
DEFINITION Rattus norvegicus clone CH230-55A15, WORKING DRAFT SEQUENCE.
ACCESSION AC114704
VERSION AC114704.4 GI:25138818
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULFILL.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 218817)
REFERENCE
Muzny,D.,Marie., Metzker,M.,Lee., Abrarzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alebrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bennahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Burch,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Ande,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinn,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Freaser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Geisregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunnaracne,F., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K.,
Harvey,I., Havlak,P., Hawes,A., Henderson,N., Hernandez,D.,
Herrandez,R., Hines,S., Hladum,S.V., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Huijck,S., Hume,C., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
LorenzLewu,L., Louisedge,H., Lozado,R.C., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mamoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundaasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parke,K.,
Pwackelemeu,O., Okwonu,G., Olarpunsaagoun,A., Pal,S., Parke,K.,
Pastermak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,

```

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Rell, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmant, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, K.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
AUTHORS
REFERENCE
TITLE
JOURNAL

Submitted (11-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 218817)
Rat Genome Sequencing Consortium.

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atlantis/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu

Project Information

Center project name: GSBS

Center clone name: CH230-55A15

Assembly program: Phrap; version 0.990329

Consensus quality: 207315 bases at least Q40

Consensus quality: 210263 bases at least Q20

Estimated insert size: 213462; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 218817: contig of 218817 bp in length.

FEATURES
source

1..218817

ORIGIN
Alignment Scores:
Pred. No.: 2,476-58 Length: 218817
Score: 1513.50 Matches: 326
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 66.26% Mismatches: 14
Query Match: 82.89% Indels: 151
DB: 2 Gaps: 4
FIGURE9 (1-361) x AC114704 (1-218817)
QY 12 Pro**GlyGlnLeuAspSerLeuSerAspLeuLeuSerArgProGlnHisGlyIle 31
122574 CCTCAGGGGAGCTGAGCTCCCTCTCAGACCTCTTCTCTCCCTCAGACAAGGAGAT 122515
QY 32 AsnLeuThr**LeuPheLeuGlyProArgCysArgAlaSerValLeuSerLeuLysAla 51
122514 AACCTCAGCTGACTGTTCTTGGGTCCTCCGCGGCGGACAGCTCTCTCCCTCAGAGCA 122455
QY 52 IleProLysCysLeuSer**GlySerLeuGlySerSerValValGlyAsnLeuGly 71
122454 ATCCCAAGTGTCTGTATGAGGCTCTTGGGAGATCTGTGTGTGGAAACCTGGGA 122395
QY 72 ThrAspAlaGlnArgLeuGlyTyrrArgValLeuProSerSerGlySerAlaAlaLeuSer 91
122394 ACAGATGACACAGAGCTGGGAGACAGATCCCTCTCTGGGCTGAGCGCTTAGC 122335
QY 92 CysSerPheProHisSerGly---PheAlaIleValArgLysCys-----LysGlyThr 108
122334 TGTTCCTTCCCAAGGCGGCGGACAGTCCGATCGGCCCAAGTGCACAGAGGCGAC 122275
QY 109 GlyMetGluTyrrAlaAlaLysPheIleLysLysArgArgLeuProSerSerArgArgGly 128
122274 GGCATGAGATGAGCGGCAAGTTCATAAAGAGCGCCCTCGCGCTCCAGCGCGCGGT 122215
QY 129 ValSerArgGlnGluIleGluArgGluValSerIleLeuArgGluIleArgHisProAsn 148
122214 GTGAGCGGTGAGAGAGATGAGCGGAGGATGAGATCTTCGCGAGATCCGCCACCCCAAC 122155
QY 149 IleIleThrLeuHisAspValPheGluAsnLysThrAspValIleLeuIleLeuGluLeu 168
122154 ATCATCAGCTCCAGATGTGTTGCAGAAACAAGACAGATGTGATCTTGAGAGCTG 122095
QY 169 ValSerGlyGlyGluLeuPheAspPheLeuLysGluLysAsp-His**GlnArgMetArg 188
122094 GTGTCCGGGGGAGACTTTTCATCTTCTGGTGGAAGAGTCACTACAGAGGATGAG 122035
QY 188 GProArgSerSerSerSerArgSerTyrThrValSerThrThrCysThrProSerAlaIse 208
122034 GCCACGCACTTCTCTCAGAGAGATCTTGACGGGTGCCACTACCTCCAGAGGATC 121975
QY 208 rArgThrLeuThr----- 212
121974 GCGCACTTTGACCTGAAGGTAGCTTGTGTGATGGCGCTGGGAGGAGGAGGAGGATTA 121915
QY 213 -----**SerArgArgThrSerCysCysTyrThr 222
121914 ACCAAGGAGACACACACACACCCCGCTCTCGCAGCCGAGAGATATATGTGTGAGAC 121855
QY 222 rSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerThr 242
1..218817

Center project name: GBW
Center clone name: CH230-SB18
----- Summary Statistics

Assembly program: Atlas;
Consensus quality: 227791 bases at least Q40
Consensus quality: 230373 bases at least Q30
Consensus quality: 232086 bases at least Q20
Estimated insert size: 237543; sum-of-contris estimation
Quality coverage: 8x in Q20 bases; sum-of-contris estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbankdraft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 244205: contig of 244205 bp in length
* 244206 244305: gap of unknown length
* 244306 245487: contig of 1182 bp in length
* 245488 245587: gap of unknown length
* 245588 246796: contig of 1209 bp in length.

FEATURES
source
Location/Qualifiers
1..246796
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-SB18"
10347..199080
/note="clone boundary
clone end:sp6
site:ECORI
end_sequence:BH334507"
complement(241125..241693)
/note="clone boundary
clone end:T7
site:ECORI
end_sequence:BH334504"

misc_feature

misc_feature
complement(241125..241693)
/note="clone boundary
clone end:T7
site:ECORI
end_sequence:BH334504"

ORIGIN

Alignment Scores:
Pred. No.: 2.78e-58 Length: 246796
Score: 1513.50 Matches: 326
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 66.26% Mismatches: 14
Query Match: 82.89% Indels: 151
DB: 2 Gaps: 4

FIGURE9 (1-361) X AC094643 (1-246796)

QY 12 Pro**GjyGlnLeuAspSerLeuSerAspLeuLeuSerArgProGlnHisGlyIle 31
DB 235301 CCGTCAGGCGAGCTGCGAATCCCTCTCAGACCTCTTCTCTCTCCCTCAGACGGGAT 235242
QY 32 AsnLeuThr**LeuPheLeuGlyProArgGlyArgAlaSerValIleuSerIleuAla 51
DB 235241 AACCTCAGTGTACGCTGTTGGGTCCCGGTGCGGCGCAGCGTCTCTCCCTCAAGCA 235182
QY 52 IleProGlyCysLeuSer**GlySerLeuGlySerSerValValAlaGlyAsnLeuGly 71
DB 235181 ATCCCGAAGTGTCTGTCTATGAGGCTCTTGGGCGAGTGTCTGTGTGGGAAACCTGGGA 235122
QY 72 ThrAspAlaGlnAgleuGlyTyrArgValLeuProSerSerGlySerAlaAlaLeuSer 91
DB 235121 ACAGATGACAGAGGCTGGGCTGACAGATCCTGCTCTCTCTGAGGCTTACG 235062
QY 92 CysSerPheProHisSerGly---PheAlaIleValArgGlyCys-----LysGlyThr 108
DB 235061 TGTCTCTTCCCGCAGCGGCGCAGTTCGCCATCTGTGGCGAAGTCCACAGAGGCGAC 235002

QY 109 GlyMetGluTyrAlaAlaLysPheIleLysLysArgAlaLeuProSerSerArgGlyIle 128
DB 235001 GGCAATGAGTACGCGGCCCAAGTCTATAAGAACGGCGCCCTCCGCTCCAGCGCGGGT 234942
QY 129 ValSerArgGluGluIleGluArgGluValSerIleLeuArgGluIleArgHisProAsn 148
DB 234941 GCGACCCCTGAGAGATGAGCGCGAGGTGAGCATCTCCCGCGAGATCCGACCCCAAC 234882
QY 149 IleIleThrLeuHisAspValPheGluAsnLysThrAspValValIleuIleLeuGluLeu 168
DB 234881 ATCATACGCTGCACGATGTGTTCCAGAACAGACAGATGTGTGCTGATCTTGGAGCTG 234822
QY 169 ValSerGlyGlyGluLeuPheAspPheLeuAlaGluLysAsp-His**GlnArgMetArg 188
DB 234821 GTGTCGGGGCGGAATTTGCACTTCTGCTGCTGAGAGAGTCACTGACAGAGATGAG 234762
QY 188 GProArgSerSerSerSerArgSerTyrThrValSerThrThrCysThrProSerAlaSe 208
DB 234761 GCGACGAGTTCCTCAAGCAGATCTGAGCGGTGCTCACTGCACTCCAAAGCGCATC 234702
QY 208 rArgThrLeuThr----- 212
DB 234701 GCGCACTTTGACCTGAGAGTAGCTTGTGTGAGTGGCGTGGGAGGAGGCGACGATA 234642
QY 213 -----SerArgArgThrSerCysStrProTh 222
DB 234641 ACCAAGGAGACACACACACACCCCGCTCTCGACGCGGAGAACATATCTGTGTGAC 234582
QY 222 rSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerArg 242
DB 234581 AAGCATGACAGCGACGCCACGATTAAGCTCATGACTTGGCATGGCGCACAGATGAG 234522
QY 242 GProValAlaSerSerArgThrSerLeuAlaArgGlnSerSerSerValArgGlyArgGly 262
DB 234521 GCGCGTAGGAGATTGAAACATCTTTGACACGCGAGGTGCGCGAGGAGCGAGGTG 234462
QY 262 sGlyHisHisProIleGly**IleLeuHisGlyLeuGlyLeuThrCysLeuAsnAsp 282
DB 234461 TGGGACACACCGGATGAGGTAGATTCTCAGAGGCTTGCTGACTGCCCTCAACATCC 234402
QY 282 oValPheHisSerPro**AspCysLysLeu**ThrThrTrpLeuGlySer**TyrVal 302
DB 234401 TGTCTTCCACACGCCCTGAGATGTAACTGTAACCTGAGCTTGAGCTGAGAGCTATG 234342
QY 302 1----- 302
DB 234341 -GGTAGTGGGGTGTGACAGAGGCCCTGCTGCTCATTTCTCCCGTCTCCCGTCT 234283
QY 302 ----- 302
DB 234282 CCCCTAAGTGTCCCAACGTGTGGGCTTATTACAGTAGTGAGGGTGTGGGCGCTAGACA 234223
QY 302 ----- 302
DB 234222 GCAGATGAGAGGGGTGAGACCCCTGCTCAGACCCCACTTATCAGATCAGTGGGCTC 234163
QY 302 ----- 302
DB 234162 TGTCTAGTACACGGGGCTTGCTGTGCTGCGATAGAGTCTGAGGTTCGTGCTGTCTGGC 234103
QY 302 ----- 302
DB 234102 TCCCGAGGTGTCAAGGCAACCTGTGTGACACTGACAGACACCTTAGATGGGCGCTTA 234043
QY 303 -----G1 303
DB 234042 TGTCTCGGGGTGAGTGTGCTGCCAAGGCCATTTGACAGACTTTTCTGCTCCCTTGCAGAGA 233983
QY 303 uHisArgArgHisHisLeuHisProValSerAla**AspGlyGlnGlyProGlnThrVal 323
DB 233982 GCATGGGCTCATCACTCAATCCTGAGTGGCTGAGATGGGCGGGGCGCTCAGACATGT 233923
QY 323 lProAlaArgGlyProGlyIleArgAlaGlyThrSerAlaAsnGlyLysHisTrpGly** 343


```

Db      233822  ACCGCTAGAGGCGCCAGGATCAGGGCTGACACTGCAAACTGGAGGCTG 233863
Oy      343    *GlutSerLeu-GlyThrLeuAspMetPro 353
Db      233862  AGAGATCTCCCTGGGAACTGGCTATGCTT 233831

RESULT 4
AC073822
LOCUS   AC073822
DEFINITION Mus musculus clone RP23-85K13, WORKING DRAFT SEQUENCE, 16 ordered
pieces.
ACCESSION AC073822
VERSION   AC073822.2
KEYWORDS  HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE    Mus musculus (house mouse)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 247196)
AUTHORS  DOE Joint Genome Institute.
TITLE     Sequencing of Mouse
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 247196)
AUTHORS  DOE Joint Genome Institute.
TITLE     Direct Submission
JOURNAL   Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT   On Jul 18, 2000 this sequence version replaced g1:8910439.
-----Genome Center
Center: Joint Genome Institute
Center code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1769765
Center clone name: RPCT-23_85K13
-----
Summary Statistics
Consensus quality: 236727 bases at least Q40
Consensus quality: 243596 bases at least Q30
Consensus quality: 244906 bases at least Q20
Estimated insert size: 243000; agarose-fp estimation
Estimated insert size: 246496; sum-of-coverage estimation
Quality coverage: 7.76 in Q20 bases; agarose-fp estimation
Quality coverage: 7.65 in Q20 bases; sum-of-coverage estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
18947: contig of 18947 bp in length
18948 19047: gap of unknown length
19048 20567: contig of 1522 bp in length
20567 20670: gap of unknown length
20670 27715: contig of 7047 bp in length
27715 27816: gap of unknown length
27816 29329: contig of 1513 bp in length
29329 29330: gap of unknown length
29330 62474: contig of 33045 bp in length
62474 62475: gap of unknown length
62475 62574: gap of unknown length
62574 66039: contig of 3465 bp in length
66039 66139: gap of unknown length
66139 91228: contig of 25089 bp in length
91228 91328: gap of unknown length
91328 91329: gap of unknown length
91329 115460: contig of 24132 bp in length
115460 115560: gap of unknown length
115560 132567: contig of 17007 bp in length
132567 132568: gap of unknown length

```

```

FEATURES
source      1..247196
             /organism="Mus musculus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10090"
             /clone="RP23-85K13"
             /clone_11b="RPCT mouse BAC library 23"

ORIGIN

Alignment Scores:
Pred. No.:      7,576-32      Length:      247196
Score:          930.00      Matches:      251
Percent Similarity: 51.35%      Conservative: 16
Best Local Similarity: 48.27%      Mismatches:  71
Query Match:    50.93%      Indels:      187
DB:             2      Gaps:      8

FIGURE9 (1-361) x AC073822 (1-247196)

Oy      16  LeuAspSerLeu---SerAspLeuLeuLeuSerArgProGlnHisGlyIleAsnLeuThr 34
Db      154290 CTGGACTCCTTGTGTCAGGAGACTCTCTCTCC---CCGACACTGAGATTAAACATAC 154346
Oy      35  **LeuPheLeuGlyProArgCysArgAlaSerValLeuSerLeuAlaIleProlys 54
Db      154347 TGACCTTCTCTGGGCTCCCATCA-gccctctct-----TCCCCAACCATCCCCAAG 154399
Oy      55  CysLeuSer**GlySerLeuGlySerSerValValAlaGlyAsnLeu-GlyThrAspAl 74
Db      154400 TGTCAATCATTTGGCTCTTTGCGCGATCTGTGTTGGGAACTGTGTGATCAAAAGC 154459
Oy      74  adLnuArgLeuGlyThrArgValAlleuProSerSerGlySerAlaIleAsnLeuCysSerPh 94
Db      154460 ATAGAGGCT-gggTACAGAGTCTTGCTTCTCCGGGCTGCGGCTTAAGTCTCTCT 154518
Oy      94  eProHisSerGly---PheAlaIleValArgLysCys-----LysGlyThrGlyMetG1 111
Db      154519 CCCCTGCAGTGGCCCAATTGTCATCTGTCGCAAGTCCAGACGAGGACGACGCGGATGGA 154578
Oy      111  uTyrAlaAlaLysPheIleLysLysArgArgLeuProSerSerArgAlaGlyValSerAr 131
Db      154579 GATGACAGCAAGTTCATCAAGAAAGCGGCTGTCATCCAGCGCGCGGTGAGGCG 154638
Oy      131  gGluGluIleGluArgGluValSerIleLeuArgGluIleArgHisProAsnIleIleTh 151
Db      154639 GGAGAGAGATCGAAGCGCGAGGTAGCATCTCGGAGATCCGCCACCCCAACATCATTAAC 154698
Oy      151  rLeuHisAspValPheGluAsnLysThrAspValValleuIleLeuGluLeuValSerG1 171
Db      154699 ACTGCATATACGGTTCGGAACACAAACAGATGTGTGTGATCTCGAGGCTGGTCCGG 154758
Oy      171  yGlyLeuLeuPheAspPheLeuAlaGlyAsp-His**GlnArgMetArgProArgS 191
Db      154759 TGGCGAGCTTTTCGCTTCCTCGCGCGAAGAGATCATTAAGCAGATAGAGCCACCA 154818
Oy      191  eSerSerSerArgSerTrpThrValSerTrpThrCysThrProSerAlaSerArgThrL 211
Db      154819 GTTCTCAACAACAAATCTTAGACGGTGTCACTACTGACCTCAAGCCCATGACACTT 154878
Oy      211  euThr----- 212

```

```

Db      154879 TGACCTGAAGTGGCGCACTTTGTGTGCGTGGAGCGCGGAGAGGCGGCGAGAGGCTA 154938
      |||||
Qy      213 -----***SerArgThrSerCysCysTyrPheSer 223
Db      154939 ACCACGAGTACACACCAATGCTGTCCCGGAGCGCGGAGAACATCATGTTCTGTGACAAGC 154998
Qy      224 MetClnProAlaHisAlaLeuSerSerSerThrLeuAlaSerArgThrGlySerArgPro 243
      |||||
Db      154999 ACCGACCGAGCGCGCGCGCATTAAGCTTCACTTGTGCGATCGCGGACAGAGATCGAGGCTG 155058
Qy      244 ValAlaSerSerArgThrSerLeuAlaArgGlnSerSerSerValArgGlyArgCysGly 263
      |||||
Db      155059 GCACGCGAGTTCAGAACATCTTTGTGACACCGCGAGTTTGGGAGAGGCGGCGAGGTGGG 155118
Qy      264 HistProIleGly***IleLeuHisGlyLeuGlyLeuThrCysLeuAlaSerProVal 283
      |||||
Db      155119 CACGCGCCAGAGGGGTGGTCTGTCACAGCCACGCTTCACTGCGCCCAACATCCCTGTC 155178
Qy      284 PheHisSerPro***AspCysIleLeu***ThrThrTyrLeuGly----- 298
      |||||
Db      155179 TCCACAGACCGCGCGAGATCGTGAACATAGAGCACTTGGGCTTGGAGGCTGACATGTGTG 155238
Qy      298 ----- 298
Db      155239 AGTGGGCTCAGCGGTGGCGCTGGCCGCTTCTCTCATACTCCCGTGTCTCCCTATTCTC 155298
Qy      298 ----- 298
Db      155299 CCGGTGCTCTTATTCTCTCTCTAAGTCAACCCACCTGAAGGCTTGAAGACAGTGTGTA 155358
Qy      298 ----- 298
Db      155359 GGGTGTGGGACACTAGACAGAGCATGAGAGGGCTGAGAGCCCGCTGCTCAAGCCCAACCT 155418
Qy      298 ----- 298
Db      155419 ATCATAGCTTCATAGAGCTGTCTCAAGTCAACGCGAGCCAGAGCTCGCTCCGCCACA 155478
Qy      298 ----- 298
Db      155479 CAGAGGCTCTGAGAGTCTACTGTCTGCTTACCGCCAGATGCTCAGAGACCTGTGTA 155538
Qy      298 ----- 298
Db      155539 CTGACACTGACACACTCTGTAGAGTGGGCGCCCTATACCTCGGTGTGAGCTGCCACGG 155598
Qy      299 -----Ser***TyrValGlnHisArgArgHisHisLeuHisPro 311
      |||||
Db      155599 CCATTCTGACCGGCTTTTCTGCGCCCTACAG-GAGCATTTGGGCTATCATCATCATCTCC 155657
Qy      312 ValSerAla***AspGlyGlnGlyProGlnThrValProAlaArgGlyProGlyIleArg 331
      |||||
Db      155658 GTGAGTGGCTGAGAGGGGCGCTTACT--AAGACTGTACTCTGTAGAGGCCAGG--ATCAG 155713
Qy      332 AlaGlyThrSerAlaLeuCys----- 338
      |||||
Db      155714 GCTGGAGACCGCTCAATATGCTAGCTGGAGATAGCTAGAGAGAACTGGGCTACTCTCT 155773
Qy      339 LysHisTyrGly***GluMetSerLeuGlyThrLeuAlaSerMetProGlyProHis 356
      |||||
Db      155774 GGCCACACCCCTAGAGCGCCCACTGTGGTGTCTAGACCCACCCCGACACACAC 155827
      |||||

```

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 1560)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Datchenko, L., Mariani, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abrahamson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Boutfield, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,

Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1560)

Strausberg, R.

Direct Submission

Submitted (10-NOV-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeff Green/Patru Kondaiah, NCI.

CDNA Library Preparation: Life Technologies, Inc.

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Galthersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nhgri.nih.gov

Akner, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Boutfield, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,

Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Startrop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, J., H. and Green, R.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 137 Row: h Column: 12

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 11968141.

Location/Qualifiers

1..1560

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="MGC:72441 IMAGE:5599435"

/tissue_type="Prostate, pool of NCI_CGAP_Pri30, 40, 41 and

NCI_CGAP_Pri29, 39, 42"

/clone_lib="NCI_CGAP_Pri50"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

1..1560

/gene="DapK1"

/db_xref="locusID:64391"

gene

ORGANISM Rattus norvegicus (Norway rat)

```

CDS
    /db_xref="RGD:621766"
    115..1461
    /codon_start=1
    /product="Death-associated like kinase"
    /protein_id="AAH62076.1"
    /db_xref="GI:38304024"
    /db_xref="LOCUSID:64391"
    /translation="MSTFRQEDVDHDEHMEGSEELSGQPAIVRKCCQKGTMEYAKFI
    KRRLPSSRRGVSRERIEVERVSLIREIRHNPNIITLHDVFENKTDVLLILEVSGGELF
    DFLAKESLTDEDEATQFLKQILDGVYHLSKRIAPDLKPEINMLDKAASPRIKLI
    DPGIARHLEAGESEFKNI FGTPEFAPEIYVPEIGLEADMSIGVITYILLGASPEFL
    GETKQETLTNISAIVNDPEEYFSELSKADFIIRLLVDPKRRMTIAQSLHSMTK
    VRRREDGAKRPERRLRARLRLEYSLSKSSMPRTSYASPPRRRVLEDAVAASQGL
    REIQRGRQCRERICALRVAAEGRACRSGSLGSDIRRTTELGRTEALTRPAQE
    EARALIGAGGLKRLCRLENNRYDALAAVAAEVQFVDLVRALEQERLQAEQVR"
    148..939
    /note="S.TKc; Region: Serine/threonine protein kinases,
    catalytic domain. Phosphotransferases of the serine or
    threonine-specific kinase subfamily. The enzymatic
    activity of these protein kinases is controlled by
    phosphorylation of specific residues in the activation
    segment of the catalytic domain, sometimes combined with
    reversible conformational changes in the C-terminal
    autoregulatory tail"
    /db_xref="CDD:c00180"

misc_feature
    /note="S.TKc; Region: Serine/threonine protein kinases,
    catalytic domain. Phosphotransferases of the serine or
    threonine-specific kinase subfamily. The enzymatic
    activity of these protein kinases is controlled by
    phosphorylation of specific residues in the activation
    segment of the catalytic domain, sometimes combined with
    reversible conformational changes in the C-terminal
    autoregulatory tail"
    /db_xref="CDD:c00180"

ORIGIN
Alignment Scores:
Pred. No.: 9,26e-31 Length: 1560
Score: 858.50 Matches: 188
Percent Similarity: 75.10% Conservative: 2
Best Local Similarity: 74.31% Mismatches: 16
Query Match: 47.02% Indels: 47
DB: 10 Gaps: 4

FIGURE9 (1-361) x BC062076 (1-1560)
QY 99 PheAlaIleValAlaGlyLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 184 TTCGCATCGTGGCGAAGTGCACACAGAGGACCCGCCAGTGAAGTACGGGGCCCAATTTC 243
QY 117 IleLysLysArgGluGlnProSerSerArgGlyValSerArgGluGluGlnIleGlnArg 136
Db 244 ATAAGAAGCCGCGCGCTGCCCTGCAGCCGCGCGGTGAGCCCTGGAGAGATGACGCGC 303
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 304 GAGGTGACATCTCGCGGAGATCCGCCACCCCAACATCAACGCTGCACGATGTGTTTC 363
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGluLeuPheAsp 176
Db 364 GAGAAACAAGACAGATGTGTCTGATCTTGAGCTGGTGTCCGCGCGCAACTTTTGAC 423
QY 177 PheLeuIleGluLysAsp-His***GlnArgMetCArgProArgSerSerSerArgse 196
Db 424 TTCTGGCTGAGAGAGTCACTGACAGAGATGAGGACAGGAGTTCCTCAAGCAGATC 483
QY 196 rTrpThrValSerThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 484 CTGGACGGGTGCCACTACCTCAGCCAGCGCATCGGCGCACTTTGACCTGAACCGGAG 543
QY 216 gThrSerGlyCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 544 AACATCAGTGTCTGGACAGACATGACGCCACGCCACATTAAGCTCATCGACTTTGGC 603
QY 236 aserArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerse 256
Db 604 ATTCGGCAGCAGATGAGCGCGGTAGCGAGTTCAAGAACATCTTTGGACGCCAGAGTTC 663
QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyLe 276
Db 664 G----- 664

RESULT 6
LOCUS RNO6971 1514 bp mRNA linear ROD 16-DEC-1998
DEFINITION Rattus norvegicus mRNA for DAP-like kinase.
ACCESSION AJ006971.1 GI:3250894
VERSION DAP-like kinase; dlx gene.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Kogel,D., Plotner,O., Landsberg,G., Christian,S. and
Scheidtmann,K.
TITLE Cloning and characterisation of dlx, a novel serine/threonine
kinase that is tightly associated with chromatin and phosphorylates
core histones
JOURNAL Oncogene 20, 2645-2654 (1998)
REFERENCE 2
AUTHORS Kogel,D.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-1998) Kogel D., Universitat Bonn, Institut fur
Genetik, Romerstr. 164, 53117 Bonn, GERMANY
REMARK 3 (bases 1 to 1514)
REFERENCE 3
AUTHORS Kogel,D.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-1998) Kogel D., Universitat Bonn, Institut fur
Genetik, Romerstr. 164, 53117 Bonn, GERMANY
FEATURES
source
    1..1514
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /db_xref="taxon:10116"
    /cell_line="SV52"
    /cell_type="rat fibroblasts transformed with SV40 large T
    antigen"
    1..1514
    /gene="dlx"
    114..1460
    /gene="dlx"
    /codon_start=1
    /product="DAP-like kinase"
    /protein_id="CAA07360.1"
    /db_xref="GI:3250895"
    /db_xref="GCA:088764"
    /db_xref="SPTREMBL:O88764"
    /translation="MSTFRQEDVDHDEHMEGSEELSGQPAIVRKCCQKGTMEYAKFI
    KRRLPSSRRGVSRERIEVERVSLIREIRHNPNIITLHDVFENKTDVLLILEVSGGELF
    DFLAKESLTDEDEATQFLKQILDGVYHLSKRIAPDLKPEINMLDKAASPRIKLI
    DPGIARHLEAGESEFKNI FGTPEFAPEIYVPEIGLEADMSIGVITYILLGASPEFL
    GETKQETLTNISAIVNDPEEYFSELSKADFIIRLLVDPKRRMTIAQSLHSMTK
    VRRREDGAKRPERRLRARLRLEYSLSKSSMPRTSYASPPRRRVLEDAVAASQGL
    REIQRGRQCRERICALRVAAEGRACRSGSLGSDIRRTTELGRTEALTRPAQE
    EARALIGAGGLKRLCRLENNRYDALAAVAAEVQFVDLVRALEQERLQAEQVR"

```

Alignment Scores:

Pred. No.:	1,11e-30	Length:	1514
Score:	856.50	Matches:	188
Percent Similarity:	75.10%	Conservative:	2
Best Local Similarity:	74.31%	Mismatches:	16
Query Match:	46.91%	Indels:	47
DB:	10	Gaps:	4

FIGURE9 (1-361) x RMO6971 (1-1514)

```

QY 99 PhealalIeValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 153 TTCCGCATCGTCCGACAGTGCAGCAGAGGCGACCCGACATGAGTACGCGGCCAAGTTC 242
QY 117 TlelylSAsArgLysLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 243 ATAAAGAGCGCGCGCTCCGTCACGCGCGCGGTGAGCCCTGAGAGATGAGCCG 302
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 303 GAGGTGACATCTCCGCGAGATCCGCCACCCCAACATCATCAGCTGCAGATGTTC 362
QY 157 GluAsnLysThrAspValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 363 GAGAACAGACAGATGTGCTGATCTTGAGCTGTGTCGCGCGCGAATCTTTTCGAC 422
QY 177 PheLeuAlaGluLysAsp-His***GlnArgMetArgProArgSerSerSerSerArgSe 196
DB 423 TTTCCGCTGAGAGAGATCTACACAGAGATGAGCCACGCTTCTTACCAATC 482
QY 186 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 483 CTGACGCGTGCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
QY 216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 543 AACATCATGTTGCTGAGACAGATGCAGCCACCCACCATTAAGCTATGACTTGGC 602
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 603 ATCCGCGACAGATGAGAGCGCGGTAGCGAGTTCAAGAACATCTTTGGACGCCAGATTC 662
QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyLe 276
DB 663 G----- 663
QY 276 uThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrT 296
DB 664 -----TCCCCCTGAGATTTGTAATGAAACCACTT 695
QY 296 rPLeuGlySer***TyrValAlaGluHisArgArgHisHisLeuHisProValSerAla***A 316
DB 696 GGCCTTGAAGCTGATATGAGATCGCGCTCATACCTTACATCCT----- 742
QY 316 spGlyGluGlyProGlnThrValProAlaArgGlyProGlyIleArgAlaGlyThrSerA 336
DB 743 -----CCTGAGCGCGGCTTCCCATTTCTCGGGGAGACCAAG 779
QY 336 laasn-----CysLysHisTTPGly***GluMet 345
DB 780 CAGAAACGCTGACAAACATCTCGGCTGTGAACTA 814

```

RESULT 7
AB007143
LOCUS AB007143 1410 bp mRNA linear ROD 25-FEB-1998
DEFINITION Mus musculus mRNA for ZIP-kinase, complete cds.
ACCESSION AB007143
VERSION AB007143.1 GI:2911153
KEYWORDS ZIP-kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS 1 (sites)
TITLE Kawai,T., Matsumoto,M., Takeda,K., Sanjo,H. and Akira,S.
JOURNAL Zip kinase, a novel serine/threonine kinase which mediates
MEDLINE apoptosis
PUBMED Mol. Cell. Biol. 18 (3), 1642-1651 (1998)
98147805
9486481
REFERENCE 2 (bases 1 to 1410)
AUTHORS Akira,S. and Kawai,T.
TITLE Direct Submision
JOURNAL Submitted (10-SEP-1997) Shizuoka Akira, Hyogo College of Medicine,
Department of Biochemistry; Mukogawa-cho 1-1, Nishinomiya, Hyogo
663, Japan (E-mail:akira@hyo-med.ac.jp, Tel:+81-798-45-6357,
Fax:+81-798-46-3164)

FEATURES
source
1..1410
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
10..1356
/codon_start=1
/product="ZIP-kinase"
/protein_id="BA024954.1"
/db_xref="GI:2911154"
/translation="MSTPRQEDVDHDEYEMGEELGSGPAIVRKQKQKMGNEYAKFI
KRLRLPSRRGVSRREIREVSLIREHPNITLHDVFNKKTVLLLELVSGELF
DLAKESITDEATOPKQILDGVHYHSKRIAPDLKPNIMVLDGHAAPRLKI
DGIARHTEAGSEFPNIGTBPFAPELVNPEIGEDMMSIGVITVILISGASPL
GETKQETLITNIAVNDPDEBFESSTSLADPRIRLIVKPKRMVTAOSLESMIX
VRRKDGAKERRRLRLRLARLESYLSKHSSEPNRTSASERSRVRVEDVAABQGL
RELQGRQCRERVCALPAAAEQREARQDSAGIGRLRLRLRELGTETALRTQAE
EARAALILGGGILKRLICRLLENRYDALAAQVAEYQFVRLVLALEOERLQAEQVR"
1410
/note="19 a nucleotides"

polya_site

ORIGIN

Alignment Scores:

Pred. No.:	1,28e-24	Length:	1410
Score:	722.00 <td>Matches:</td> <td>163 </td>	Matches:	163
Percent Similarity:	73.82% <td>Conservative:</td> <td>9 </td>	Conservative:	9
Best Local Similarity:	69.96% <td>Mismatches:</td> <td>21 </td>	Mismatches:	21
Query Match:	39.54% <td>Indels:</td> <td>40 </td>	Indels:	40
DB:	10 <td>Gaps:</td> <td>5 </td>	Gaps:	5

FIGURE9 (1-361) x AB007143 (1-1410)

```

QY 99 PhealalIeValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 79 TTTCGATCGTCCGACAGTGCAGCAGAGGCGACGCGCATGAGTACGAGCCAGTTC 138
QY 117 TlelylSAsArgLysLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 139 ATCAAGAGAGCGCGCTGTCATCCAGCCGCGCGGTGAGAGCGGAGAGATCGAACGC 198
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 199 GAGGTGAGATCTCCGCGAGATCCGCCACCCCAACATCATACCTGCTGAGAGCTGTC 258
QY 157 GluAsnLysThrAspValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 259 GAGAACAGACAGATGTGCTGATCTGAGCTGTGTCGCGGTGCGAGACTTTTCGAC 318
QY 177 PheLeuAlaGluLysAsp-His***GlnArgMetArgProArgSerSerSerSerArgSe 196
DB 319 TTTCGCGCGAAGAGATCATGAGAGATGAGGCGACGCACTTCTCAACAAATATC 378
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 379 CTAGACGGTGTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
QY 216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236

```

DB 439 AACATCATGTTGCTGACAGACGACGAGCCGCCGATTAAGCTCATGACTTTGGC 498
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 499 ATCGCGACAGGATCGAGGCTGCGAGGAGTTCAAGAACATCTTTGGCA----- 547
QY 256 rSerValArgGlyArgCysGlyHisHisPro-IleGly**IleLeuHisGlyLeuGlyL 276
DB 548 -----CACCCGAGTT----- 557
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
DB 558 -----TGT-----CGCCCCGAGATCGTGAACATATGAGCCACTT 591
QY 296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
DB 592 GGCTTGAGGCTGACATGAGCATGAGCATGCGTCACTACCTACCTCTGAG----- 644
QY 316 spGlyGlnGlyProGlnThrValProAlaArgGly 327
DB 645 -----CGAGCGCTCCCATCTCTGGG 665
RESULT 8
AR076190 1429 bp DNA linear PAT 30-AUG-2000
LOCUS Sequence 4 from patent US 5958748.
DEFINITION AR076190
ACCESSION AR076190
VERSION AR076190.1 GI:10002936
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1429)
AUTHORS Akira,S. and Kawai,T.
TITLE DNA coding for serine/threonine kinase
JOURNAL Patent: US 5958748-A 4 28-SEP-1999;
FEATURES
Location/Qualifiers
1..1429
source /organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.3e-24 Length: 1429
Score: 722.00 Matches: 163
Percent Similarity: 73.82% Conservative: 9
Best Local Similarity: 69.96% Mismatches: 21
Query Match: 39.54% Indels: 40
DB: 6 Gaps: 5
FIGURES (1-361) x AR076190 (1-1429)
QY 99 PheAlaIleValAlaArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 79 TTGGCATCTGTCGCGCAAGTGCAGACAGAGGACGCGGATGAGTATGACGCAAGTTTC 138
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 139 ATCAAGAAAGCGGCGCTCCATCCAGCGCGGCGGTGTGAGCCGGAGAGATCGAACGC 198
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 199 GAGGTGAGCATCTCGCGCGAGATCCGCCAACCACTATACATGATGATGCGTTC 258
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 259 GAGAAACAAGACAATGTGTGCTGATCTGAGCTGTGTCCGATCGCGAGCTTTTCGAC 318
QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSerSerArgSe 196
DB 319 TTCTGCGCCAGAAAGATGATGATGACGAGATGAGGCCCAAGATTCTCTCAACAATC 378
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216

DB 379 CTAGACGCTGTCTACTTACTGCTACCTCCAGCGCATCGACACTTGAAGCCCGAG 438
QY 216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 439 AACATCATGTTGCTGACAGACGACGAGCCGCCGATTAAGCTCATGACTTTGGC 498
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 499 ATCGCGACAGGATCGAGGCTGCGAGGAGTTCAAGAACATCTTTGGCA----- 547
QY 256 rSerValArgGlyArgCysGlyHisHisPro-IleGly**IleLeuHisGlyLeuGlyL 276
DB 548 -----CACCCGAGTT----- 557
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
DB 558 -----TGT-----CGCCCCGAGATCGTGAACATATGAGCCACTT 591
QY 296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
DB 592 GGCTTGAGGCTGACATGAGCATGAGCATGCGTCACTACCTACCTCTGAG----- 644
QY 316 spGlyGlnGlyProGlnThrValProAlaArgGly 327
DB 645 -----CGAGCGCTCCCATCTCTGGG 665
RESULT 9
AR124103 1429 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 4 from patent US 6171841.
DEFINITION AR124103
ACCESSION AR124103
VERSION AR124103.1 GI:14109464
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1429)
AUTHORS Akira,S. and Kawai,T.
TITLE DNA coding for serine/threonine kinase
JOURNAL Patent: US 6171841-A 4 09-JAN-2001;
FEATURES
Location/Qualifiers
1..1429
source /organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.3e-24 Length: 1429
Score: 722.00 Matches: 163
Percent Similarity: 73.82% Conservative: 9
Best Local Similarity: 69.96% Mismatches: 21
Query Match: 39.54% Indels: 40
DB: 6 Gaps: 5
FIGURES (1-361) x AR124103 (1-1429)
QY 99 PheAlaIleValAlaArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 79 TTGGCATCTGTCGCGCAAGTGCAGACAGAGGACGCGGATGAGTATGACGCAAGTTTC 138
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 139 ATCAAGAAAGCGGCGCTCCATCCAGCGCGGCGGTGTGAGCCGGAGAGATCGAACGC 198
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 199 GAGGTGAGCATCTCGCGCGAGATCCGCCAACCACTATACATGATGATGCGTTC 258
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 259 GAGAAACAAGACAATGTGTGCTGATCTGAGCTGTGTCCGATCGCGAGCTTTTCGAC 318

CDS

46..1410
/gene="ZIPK"
/codon_start=1
/product="ZIP kinase"
/protein_id="BA81746.1"
/db_xref="GI:5162884"
/translation="MSTFROEDVEDHYEMWCEELSGQFAIVKCKQKGTGKYEAKFI
KRRLLSSRRGVREIREVENLITREIRHNITLHDIPENKIDVLLILEVSGGLF
DPLAKESLITREDAEITQFLOKIDGVHITSKRIIAHDIKPNIMLDDKVPNRIKLI
DPTIAKIEAKNEFKNIFGTPPEVPAEIVNVEPLGADMSIGITTYLLISGASPEL
GETRQETLNTISAVNTDFDEYFNTSELMKDFIRLIVKPRRTTIKQSLPHSMIK
AIRRNVRGDSGRKPERRLKTRLEKVTIKSHSSLPNNSTADERSKVLSEAA
AEEGLREIQSRRLCHEDVLAALAIYEKEMWRESDSLIGDLRELRQELKTEALK
ROAOKERKALGTSGIKRFRSLRNRYEALAKQAVSEWRFVDLVRALOEKLOQVE
GGR"

2055
/gene="ZIPK"
/note="24 a nucleotides"

polyA_site

ORIGIN

Alignment Scores:

Pred. No.:	2,21e-20	Length:	2055
Score:	632.00	Matches:	155
Percent Similarity:	62.69%	Conservative:	13
Best Local Similarity:	57.84%	Mismatches:	46
Query Match:	34.61%	Indels:	54
		Gaps:	9

FIGURE9 (1-361) x AB022341 (1-2055)

```

QY 99 PheAlaIleValAArgLysCys-----LysGLYThrGLYMetGLUTyrAlaAlaLysPhe 116
Db 115 TTTCCGATCGCGCGAAGTGCAGAGCGGCGAGGCGGCGGCGAGAGTACGCGCGCAAGTTC 174
QY 117 IleLysLysAArgAArgLeuProSerSerAArgGLYValSerAArgLugLugLugLugAArg 136
Db 175 ATCAAGAAAGCCGCCCTGTATCATCCAGCGGCGTGGGTGAGCGCGGAGGAGATCGAGCGG 224
QY 137 GluValSerLLeuAArgGluLLeAArgHisProAsnLLeLLeThrLeuHisAspValPhe 156
Db 235 GAGGTGAACATCCCGCGGAGATCCGCGACCCCAACATCATCACCCCGCAGCAGATCTTC 294
QY 157 GluAsnLysThrAspValAlaLeuLLeLLeuGluLeuValSerGLYGLYGLULeuPheAsp 176
Db 295 GAGAACAGAGGACGAGTGTCTCATCTCGAGCTGTCTTGGCGGAGGAGCTTTTGAC 354
QY 177 PheLeuAlaGluLysAspHis**GlnAArgMetAArgProAArgSerSerSerAArgSe 196
Db 355 TTCCCTGGCGGAGAGAGATCGCTGACGAGAGCGAGGCCACCCAGTTCTCTCAAGCAGATC 414
QY 196 TTTPThrValSerThrThrCysThrProSerAlaSerAArgThrLeuThr**SerAArgAr 216
Db 415 CTGACCGCGGTTCACTACCTGCACTTAAGGCAATCCGACATTCCTGACCTGMAAGCGGAA 474
QY 216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 475 AACATCATAGTGTCTGGAAGAAGAGAGTGCACCAACCAAGATCAAGCTCATCTGCTGGC 534
QY 236 aSerAArgThrGLYSerAArgProValAlaSerSerAArgThrSerLeuAlaArgGlnSerSe 256
Db 535 ATCGCGGCAAGATCGAGCGGCGGAGACAGATTCAAGACAA-----TC 576
QY 256 rSerVal_AArgGLYArgCysGLYHisHisProLLeGLY**LLeLeuHisGLYLeuGlyL 276
Db 577 TTTCGACACCCCGGAGTTGTGGC----- 599
QY 276 eutHrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
Db 600 -----CCAGAGATTGTGACATCTGAGCCGCTG 627
QY 296 rPLeuGLYSer**TyrValGluHisAArgAArgHisHisLeuHisProValSerAla**A 316
Db 628 GGCCTGGAGCGGAGCATGTGAGCATCGGTGTCTCATCTACCTATATCT----- 674
  
```

QY 316 spGLYGLNGLYProGLN-----ThrValProAlaAArgGLYProGLYLLeAArgAlaG 333
Db 675 -----CTCGAGCGGTGCATCCCGCTCTGCGCGAGAC-----CAACGAG 714

QY 333 lYthrSerAlaAsnCysLysHisTTPGLY**GluMetSerLeu-----GLYThrLeuA 351
Db 715 GAGACGCTCAC-----CAACATCTCACCCGCTGAACATCAAGCTTCGACGAGGAGTACTTC 768

QY 351 spMetProGLYProHisGLN 357
Db 769 AGCAA-----CACCG 779

RESULT 13
LOCUS BD127287 2079 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127287
VERSION BD127287.1 GI:2322232
KEYWORDS UP 2002017375-A/2718.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2079)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Kakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
REFERENCE
BD127287
Patent: JP 2002017375-A 2718 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN UP 2002017375-A/2718
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TARAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUTSU OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
FT CDS Location/Qualifiers
source 1..2079
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	2,24e-20	Length:	2079
Score:	632.00 <td>Matches:</td> <td>155 </td>	Matches:	155
Percent Similarity:	62.69% <td>Conservative:</td> <td>13 </td>	Conservative:	13
Best Local Similarity:	57.84% <td>Mismatches:</td> <td>46 </td>	Mismatches:	46
Query Match:	34.61% <td>Indels:</td> <td>54 </td>	Indels:	54
		Gaps:	9

FIGURE9 (1-361) x BD127287 (1-2079)

```

QY 99 PheAlaIleValAArgLysCys-----LysGLYThrGLYMetGLUTyrAlaAlaLysPhe 116
Db 139 TTTCCGATGTGCGGAAGTGCAGAGGCGGCGGCGGCGGCGGCGAGTACGCGCGCAAGTTC 198
QY 117 IleLysLysAArgAArgLeuProSerSerAArgGLYValSerAArgLugLugLugLugAArg 136
Db 199 ATCAAGAAAGCCGCCCTGTATCATCCAGCGGCGTGGGTGAGCGGCGGAGAGATCGAGCGG 258
QY 137 GluValSerLLeuAArgGluLLeAArgHisProAsnLLeLLeThrLeuHisAspValPhe 156
  
```



```

Db      652 GGCCTGAGGCGGACATGTGAGACATGCGTGTATCATCACTATATCTCT----- 698
Qy      316 spgIyngIyProGln-----ThyValProAlaArgGlyProGlyIleArgAlaG 333
Db      699 -----CCTGACCGGTCGATCCCGTCTCTCGGCGAGAC-----CAAGCAG 738
Qy      333 lYThSerAlaAsnCysLysHisTrpGly**GIuMetSerLeu-----GIYThLeuA 351
Db      739 GAGAGCGCTCAC-----CAACATCTCAGCCGCTGACTACGACTTCGACGAGGACTTTC 792
Qy      351 spMetProGlyProHisGln 357
Db      793 AGCAA-----CACGAG 803

RESULT 15
LOCUS      BD176607          2105 bp      DNA      linear      PAT 18-MAR-2003
DEFINITION Method of molecular diagnosis of chronic myelogenous leukemia.
ACCESSION  BD176607
VERSION     BD176607.1 GI:29122317
KEYWORDS    WO 02070747-A/9
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 2105)
AUTHORS     Mamu, H.
TITLE        Method of molecular diagnosis of chronic myelogenous leukemia
JOURNAL     Patent: WO 02070747-A 9 12-SEP-2002;
            FUJISAWA PHARMACEUTICAL CO LTD, HIROYUKI YAMMO
COMMENT      OS Homo sapiens (human)
            PN WO 02070747-A/9
            PD 12-SEP-2002
            PE 01-MAR-2002 WO 2002JP001901
            PR 01-MAR-2001 JP 01P 056438
            PI HIROYUKI YAMMO
            PC C12Q1/68, C12Q1/02, G01N33/15, C12N15/09, A61P35/02 CC
            molecular diagnosis of chronic myelogenous leukemia FH
            Location/Qualifiers
              FT CDS             (94)..(1458).
               source          1..2105
                           /organism="Homo sapiens"
                           /mol_type="genomic DNA"
                           /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      2,27e-20      Length:      2105
Score:          632.00      Matches:      155
Percent Similarity: 62.69%      Conservative: 13
Best Local Similarity: 57.84%      Mismatches: 46
Query Match:    34.61%      Indels:      54
DB:              Gaps:      9

FIGURE9 (1-361) x BD176607 (1-2105)
Qy      99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db      163 TTTCGATCGTGCAGAACTGCCGAGAGGACGCGGCAAGTACGACGCCAAGTTC 222
Qy      117 lIeLysIysArgArgLeuProSerSerArgGlyValSerArgGluGluArg 136
Db      223 ATCAAGAGCGCGCGCTCTCTCATCCAGCGCGCTGGGTGAGCCGGAGAGATCGAGCGG 282
Qy      137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db      283 GAGGTGAACATCTCGCGGAGATCCGGACCCCAACATCATCACCTCGACGACATCTTC 342
Qy      157 GluAsnIysThrAspValValLeuIleGluGluLeuValSerGlyGlyGluLeuPheAsp 176
Db      343 GAGAACAAAGACGAGAGTGTCTCTCTGAGAGCTGGGAGGAGCTTTTGAC 402

```

```

Qy      177 PheLeuAlaGluLysAspHis**GIuArgMetArgProArgSerSerSerArgSe 196
Db      403 TTTCCTGGGCGGAGAAAGATGCTGTACCGGAGACAGACGCCACCGCTCTCAACAGATC 462
Qy      196 lTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
Db      463 CTGACGGCGGTTCACCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 522
Qy      216 gThSerCysCysTrpThrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db      523 AACATCACTGCTGCTGAGACAGACAGTCCCAACCCACAGAACAGTCACTGATCTTGCGG 582
Qy      236 AserArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db      583 ATCGGACACAGATCGAGGCGGAGACGAGTTCAGAAACA-----TC 624
Qy      256 rSerValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyL 276
Db      625 TTGGGACCCCGGAGTTGTGGC----- 647
Qy      276 euThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
Db      648 -----CCAGAGATTGTGAACATAAGAGCCGCTG 675
Qy      296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
Db      676 GGCCTGAGGCGGACATGTGAGCATCGTGTATCATCACTATATCT----- 722
Qy      316 spgIyngIyProGln-----ThyValProAlaArgGlyProGlyIleArgAlaG 333
Db      723 -----CCTGACCGGTCGATCCCGTCTCTCGGCGAGAC-----CAAGCAG 762
Qy      333 lYThSerAlaAsnCysLysHisTrpGly**GIuMetSerLeu-----GIYThLeuA 351
Db      763 GAGAGCTCAC-----CAACATCTCAGCCGTAACATAAGACTTCGACGAGGAGTACTTC 816
Qy      351 spMetProGlyProHisGln 357
Db      817 AGCAA-----CACGAG 827

```

Search completed: July 1, 2004, 03:41:50
Job time : 6338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2004, 18:49:15 ; Search time 3354 Seconds
(without alignments)
9731.474 Million cell updates/sec

Title: US-10-083-641a-6
Perfect score: 1093
Sequence: 1 gnatgnatcggttaac.....ccacaaangtagaccatnc 1093

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbu:*
5: em_estbv:*
6: em_estbp:*
7: em_estbr:*
8: em_estbs:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estgm:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	363	33.2	666 13	BU670814 NISC_1r01
2	343	31.4	637 14	CA334157 NISC_1s06
3	267	24.4	597 13	BU671103 NISC_1r05
4	217	19.9	417 14	CB766821 AMGNNTC:IS

5	193	17.7	565 10	BF547702 UI-R-A0-b
6	178	16.3	448 14	CB746201 AMGNNTC:N
7	171	15.6	412 14	CB801518 AMGNNTC:S
8	150	12.7	440 14	CB745794 AMGNNTC:M
9	131	12.0	455 14	CB741392 AMGNNTC:N
10	126	11.5	410 14	CB803207 AMGNNTC:M
11	123	11.3	411 14	BG663786 DRAMNF02
12	103	9.4	451 14	CB787038 AMGNNTC:M
13	44	4.0	136 9	AA965990 ua38h03.r
14	44	4.0	216 9	AA027503 mi04e11.r
15	44	4.0	235 9	AA027581 mi06c09.r
16	44	4.0	316 9	AA864020 vx88a02.r
17	44	4.0	331 9	AA027379 mi05c03.r
18	44	4.0	340 9	AA013814 mh06f01.r
19	44	4.0	343 10	BE654598 UI-M-A71-
20	44	4.0	354 9	AA982305 ua53b01.r
21	44	4.0	376 9	AA023436 mh72f07.r
22	44	4.0	385 9	AA024296 mh92a05.r
23	44	4.0	395 14	W75433 me50d04.r1
24	44	4.0	397 9	AA020495 mh61g04.r
25	44	4.0	399 9	AA013793 mh13e04.r
26	44	4.0	400 9	AL363655 AL363655
27	44	4.0	405 10	BB853739 ux20d04.y
28	44	4.0	422 10	BB849679 BB849679
29	44	4.0	423 9	AA050632 mj18e05.r
30	44	4.0	423 14	W85231 mf52b04.r1
31	44	4.0	428 13	BY040500 BY040500
32	44	4.0	437 10	BF147326 uw68h08.y
33	44	4.0	439 14	W77571 me67e12.r1
34	44	4.0	443 9	AA475990 vlt25g12.r
35	44	4.0	446 9	AA023330 mh71b08.r
36	44	4.0	446 9	AA023330 mh71b08.r
37	44	4.0	449 10	BF012286 ux86b04.r
38	44	4.0	449 10	BF012286 ux86b04.r
39	44	4.0	454 9	AA023742 mh81b03.r
40	44	4.0	456 9	AA017794 mh47g09.r
41	44	4.0	457 9	AA015164 mh36e06.r
42	44	4.0	459 9	AA016932 mh37g08.r
43	44	4.0	465 9	AA017803 mh47a03.r
44	44	4.0	475 9	AA048970 mj50c05.r
45	44	4.0	482 10	BB629354 uo37a01.y
			483 9	AA467508 ve01c12.x

ALIGNMENTS

RESULT 1
BU670814 666 bp mRNA linear EST 01-OCT-2002
LOCUS NISC_1r01e11.y1 NCI CGAP Pr49 Rattus norvegicus cDNA clone
DEFINITION IMAGE:5598044 5', mRNA sequence.
ACCESSION BU670814.1 GI:23398790
VERSION BU670814
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 666)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation:
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov

QY 587 CAGGACATCTGGACGGGTGTCCACTACCTGCACTCCAGCGCATGCGCACTTTGACT 646
Db 474 CAGGACATCTGGACGGGTGTCCACTACCTGCACTCCAGCGCATGCGCACTTTGACT 533
QY 647 GAAGCCGGAACATCATCTGTTGCTGACAGACATGACGAGCCGACGATTAAGTCAT 706
Db 534 GAAGCCGGAACATCATCTGTTGCTGACAGACATGACGAGCCGACGATTAAGTCAT 593
QY 707 GCACTTTGGCATGCGCAGACAGATCGAGCGCGGTAGCGAGTTCA 750
Db 594 GCACTTTGGCATGCGCAGACAGATCGAGCGCGGTAGCGAGTTCA 637
RESULT 3
LOCUS BU671103 597 bp mRNA linear EST 01-OCT-2002
DEFINITION MISC_l1r05b04.y1 NCI_CGAP_P149 Rattus norvegicus cDNA clone
IMAGE:5598270 5', mRNA sequence.
BU671103
BU671103.1 GI:23399332
EST
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 597)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncigap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: rgsb@r-xmail.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
info@image.llnl.gov
Plate: LHAM12385 row: C column: 7
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..597
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:5598270"
/sex="male"
/tissue_type="ventral prostate, pool of 3-, 5-, and 7-days
post-castration"
/dev_stage="adult, 11 week"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP P149"
/note="Organ: prostate; Vector: pCMV-SORT6.1; Site:1:
NCI1; Site:2: SCORV; Cloned unidirectionally. Primer:
Oligo dt. Pool of 3 primary libraries: NCI_CGAP_P130
(ventral prostate from 11 wk male, 3 days
post-castration, average insert size 2 kb), NCI_CGAP_P140
(ventral prostate from 11 wk male, 5 days
post-castration, average insert size 1.6 kb) and
NCI_CGAP_P141 (ventral prostate from 11 wk male, 7 days
post-castration, average insert size 2.5 kb). Constructed
by Life Technologies/Invitrogen. Note: this is a NCI_CGAP
Library."

ORIGIN
Query Match 24.4%; Score 267; DB 13; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.3e-120; Indels 0; Gaps 0;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 288 CAGGCGCAGTTGCGCATGTCGCAAGTCCAGACAGAGGCGCACCGGCGATGAGTACG 347
|||||

Db 110 CAGCGGCCAGTTGCCATCTGTCGCAAGTCCAGACAGAGGCGCACCGGCGATGAGTACG 169
QY 348 GGCCAAAGTTCAATMAAGAGCGCGCTGCTCCAGCGCGCGGTGTGAGCCGTGAGA 407
Db 170 GGCCAAAGTTCAATMAAGAGCGCGCTGCTCCAGCGCGCGGTGTGAGCCGTGAGA 229
QY 408 GATCCAGCGCGGAGTGTGAGCATCTCCGCGAGATCCGCCACCCCAATCATACGCTGCA 467
Db 230 GATCCAGCGCGGAGTGTGAGCATCTCCGCGAGATCCGCCACCCCAATCATACGCTGCA 289
QY 468 CGATGTTCGGAACAACAAGATGTGTGCTGATCTTGGAGCTGTGTCCGCGCGGA 527
Db 290 CGATGTTCGGAACAACAAGATGTGTGCTGATCTTGGAGCTGTGTCCGCGCGGA 349
QY 528 ACTTTTCGACTTCTGCGCTGAGAGGA 554
Db 350 ACTTTTCGACTTCTGCGCTGAGAGGA 376
RESULT 4
LOCUS CB766821 417 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNMNC:SRPB2-00314-C4-A srpb2 (10220) Rattus norvegicus cDNA clone
srpb2-00314-c4 5', mRNA sequence.
CB766821
CB766821.1 GI:29855212
EST
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 417)
REFERENCE Amgen EST Program.
AUTHORS Amgen Rat EST Program
TITLE Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00314 row: C column: 4.
Location/Qualifiers
1..417
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srpb2-00314-c4"
/tissue_type="prostate tissue"
/note="Vector: pSPORT1; Site:1: SalI; Site:2: NotI; rat
prostate normalized double selected poly(A+) mRNA size
fraction > 1 kb"

ORIGIN
Query Match 19.9%; Score 217; DB 14; Length 417;
Best Local Similarity 99.6%; Pred. No. 6.5e-96;
Matches 267; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 649 AGCCGGAACATCATCTGTTGCTGACAGACATGACGAGCCGACGATTAAGTCATCG 708
Db 105 AGCCGGAACATCATCTGTTGCTGACAGACATGACGAGCCGACGATTAAGTCATCG 164
QY 709 ACTTTGGCATGCGCAGACAGATCGAGCGCGGTAGCGAGTTCAAGAAATCTTTGGCAGCG 768
Db 165 ACTTTGGCATGCGCAGACAGATCGAGCGCGGTAGCGAGTTCAAGAAATCTTTGGCAGCG 224
QY 769 CAGAGTTGCTGCGTAGAGGCGCAGGTGTGGGACCAACCCGATAGGTTAGTGGACGCG 828
Db 225 CAGAGTTGCTGCGTAGAGGCGCAGGTGTGGGACCAACCCGATAGGTTAGTGGACGCG 284
QY 829 CTTGGCTGACCTGCTCAACATCTGTCTTCCACAGAGCCCTGAGATTGAACTATGA 888
Db 285 CTTGGCTGACCTGCTCAACATCTGTCTTCCACAGAGCCCTGAGATTGAACTATGA 344
|||||

Qy 889 ACCACTGGCTTGGAGCTGATATGTGG 916
Db 345 ACCACTGGCTTGGAGCTGATATGTGG 372

RESULT 5
LOCUS BF547702 565 bp mRNA linear EST 11-DEC-2000
DEFINITION U1-R-A0-Df-a-02-0-UI.r1 U1-R-A0 Rattus norvegicus cDNA clone
ACCESSION BF547702
VERSION BF547702
KEYWORDS BF547702.1 GI:11638809
SOURCE EST.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 565)
REFERENCE 1
AUTHORS Ronald M.F., Lennon G. and Soares M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contract: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBR, Iowa City, IA 52242, USA
Tel.: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Clones will be available through Research Genetics (www.rsgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LNU (info@image.lnu.gov). IMAGE ID=1763507
Seq primer: M13 Forward.
Location/Qualifiers

FEATURES

source

1..565
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="U1-R-A0-Df-a-02-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pTZ19-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI. This library
consists of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung,
brain, liver, kidney, heart, spleen, ovary, and muscle.
The tag is a string of 3-5 nucleotides present between the
Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture."

ORIGIN

Query Match 17.7%; Score 193; DB 10; Length 565;
Best Local Similarity 100.0%; Pred. No. 5,4e-84;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 362 AGAAGCGCGCGCTCCGTCGAGCGCGCGGTGAGAGGAGATGACCGGAGG 421
Db 233 AGAAGCGCGCGCTCCGTCGAGCGCGCGGTGAGAGGAGATGAGCGGAGG 292
Qy 422 TGACATCTCTGGGAGATCGCCACCCCAACATCATCAGCTGACGATGTGTGAGA 481
Db 293 TGACATCTCTGGGAGATCGCCACCCCAACATCATCAGCTGAGATGTGTGAGA 352
Qy 482 ACAAGACGATGTGTGTGATCTTGAAGCTGTGTCTCGGCGGAGCACTTTGACTTTC 541

Db 353 ACAAGACGATGTGTGTGATCTTGAAGCTGTGTCTCGGCGGAGCACTTTGACTTTC 412
Qy 542 TGGCTGAGAGAGA 554
Db 413 TGGCTGAGAGAGA 425

RESULT 6
LOCUS CB746201 448 bp mRNA linear EST 11-APR-2003
DEFINITION AMGNNUC:SRPB2-00315-F4-A srbp2 (10220) Rattus norvegicus
ACCESSION CB746201
VERSION CB746201
KEYWORDS CB746201.1 GI:29813503
SOURCE EST.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 448)
REFERENCE 1
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00055 row: h column: 4.
Location/Qualifiers

FEATURES

source

1..448
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srpb2-00315-F4-A srbp2 (10220) Rattus norvegicus"
/note="Vector: pSPOR1. Site 1: SalI; Site 2: NotI; W Rat
pituitary adult female wistar rat avg insert size 2.1 kb"

ORIGIN

Query Match 16.3%; Score 178; DB 14; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.3e-76;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 288 CAGCGCGCATCTCCGATCTGCGCAATGCCAGACAGAGGACCGGCATGAGTACG 347
Db 242 CAGCGCGCATCTCCGATCTGCGCAATGCCAGACAGAGGACCGGCATGAGTACG 301
Qy 348 GGCAGATTCAATAAGAGCGCGCGCTCCGTCAGCGCGCGGTGTGACCGGTGAGA 407
Db 302 GGCAGATTCAATAAGAGCGCGCGCTCCGTCAGCGCGCGGTGTGAGCGGTGAGA 361
Qy 408 GATCGAGCGCGAGGTGACATCTTGGCGGAGATCCGCCAACATCATCAGCTG 465
Db 362 GATCGAGCGCGAGGTGACATCTTGGCGGAGATCCGCCAACATCATCAGCTG 419

RESULT 7
LOCUS CB801518 412 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNNUC:SRPB2-00315-F4-A srbp2 (10220) Rattus norvegicus cDNA clone
ACCESSION CB801518
VERSION CB801518
KEYWORDS CB801518.1 GI:29914858
SOURCE EST.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 412)
REFERENCE 1
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program

JOURNAL
COMMENT

Unpublished (2003)
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00315 row: f column: 4.
Location/Qualifiers

FEATURES
source

1..412
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srpb2-00315-f4"
/tissue_type="prostate tissue"
/clone_id="srpb2 (10220)"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; rat prostate normalized double selected poly(A+) mRNA size fraction > 1 kb"

ORIGIN

Query Match 15.6%; Score 171; DB 14; Length 412;
Best Local Similarity 100.0%; Pred. No. 3.8e-73;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 AGCCGGGAACATCATCTTCTGACACAGATGACGACCCCAAGCATTAACTCATG 708
DB 105 AGCCGGGAACATCATCTTCTGACACAGATGACGACCCCAAGCATTAACTCATG 164
QY 709 ACTTGGCATCGCGCAGAGATGAGGCGCGTAGCGAGTTCAAGAACATCTTGGGACCG 768
DB 165 ACTTGGCATCGCGCAGAGATGAGGCGCGTAGCGAGTTCAAGAACATCTTGGGACCG 224
QY 769 CAGAGTTCTGCTGAGGCGGAGGTGTGGGACACCCCGATTAGGTAGATT 819
DB 225 CAGAGTTCTGCTGAGGCGGAGGTGTGGGACACCCCGATTAGGTAGATT 275

RESULT 8
LOCUS CB749794 440 bp mRNA linear EST 11-APR-2003
DEFINITION AMGNNUC:NRW34-00018-H10-A mrp4 (10380) Rattus norvegicus cDNA

ACCESSION CB749794
VERSION CB749794.1 GI:29817096
KEYWORDS EST
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 440)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00018 row: h column: 10.
Location/Qualifiers

FEATURES
source

1..440
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrp4-00018-h10"
/tissue_type="placenta embryo"
/clone_id="mrp4 (10380)"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; placenta embryo day 17"

ORIGIN

Query Match 13.7%; Score 150; DB 14; Length 440;
Best Local Similarity 100.0%; Pred. No. 9.6e-63;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 CAGCGCCAGTTCCGATCGTGGCGAAGTGCAGAGAAAGGACCGGATGAGTACGC 347
DB 291 CAGCGCCAGTTCCGATCGTGGCGAAGTGCAGAGAAAGGACCGGATGAGTACGC 350
QY 348 GAGCAAGTTCAATAAAGAGGCGGCGCTGCGGTCAGCGCGCGGTGTAGCCGTGAGCA 407
DB 351 GAGCAAGTTCAATAAAGAGGCGGCGCTGCGGTCAGCGCGCGGTGTAGCCGTGAGCA 410
QY 408 GATCAGCGCGAGGTGAGCATCTGCGCGA 437
DB 411 GATCAGCGCGAGGTGAGCATCTGCGCGA 440

RESULT 9
LOCUS CB741392 455 bp mRNA linear EST 11-APR-2003
DEFINITION AMGNNUC:NRW33-00012-F6-A white adipose tiss (10469) Rattus

ACCESSION CB741392
VERSION CB741392.1 GI:29808681
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 455)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00012 row: f column: 6.
Location/Qualifiers

FEATURES
source

1..455
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrw33-00012-f6"
/tissue_type="adipose tiss"
/clone_id="white adipose tiss (10469)"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; white adipose tiss adult female Wistar rats, >250 grams"

ORIGIN

Query Match 12.0%; Score 131; DB 14; Length 455;
Best Local Similarity 99.6%; Pred. No. 2.5e-53;
Matches 251; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 426 CATCTGGCGAGATCCGCCACCCACATCATACGCTGACGATGTTGAGAACAA 485
DB 184 CATCTGGCGAGATCCGCCACCCACATCATACGCTGACGATGTTGAGAACAA 243
QY 486 GACAGATGTGTGCTGATCTTGAGAGCTGTGTCTCGGCGGCGAACTTTTGCATTTCGGC 545
DB 244 GACAGATGTGTGCTGATCTTGAGAGCTGTGTCTCGGCGGCGAACTTTTGCATTTCGGC 303
QY 546 TGAGAAAGA-TCATGACAGAGATGAGGCGACGAGTCTCTCAAGCAGATCTGACGG 604
DB 304 TGAGAAAGATCATGACAGAGATGAGGCGACGAGTCTCTCAAGCAGATCTGACGG 363
QY 605 TGTCCACTACTGCACTCCAGAGCCATGCGCACTTGTGACTGAGCCGGAACATCAT 664
DB 364 TGTCCACTACTGCACTCCAGAGCCATGCGCACTTGTGACTGAGCCGGAACATCAT 423
QY 665 GTTGCTGAGCAA 676
DB 424 GTTGCTGAGCAA 435

RESULT 10
CB803207 410 bp mRNA linear EST 16-MAY-2003
LOCUS CB803207
DEFINITION AMGNNUC:MRBE3-00033-D9-A placenta embryo D17 (10379) Rattus
norvegicus cDNA clone mrep3-00033-d9 5', mRNA sequence.
ACCESSION CB803207 GI:29918174
VERSION CB803207.1
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 410)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00033 row: d column: 9.
FEATURES
source Location/Qualifiers
1..410
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrep3-00033-d9"
/issue_type="placenta embryo"
/clone_id="placenta embryo D17 (10379)"
/note="Vector: pSPORT1; placenta embryo D17"
ORIGIN
Query Match 11.5%; Score 126; DB 14; Length 410;
Best Local Similarity 100.0%; Pred. No. 7.1e-51;
Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 384 CCGGCGGGGTGTGCGCGGTGAGATCGAGCGGAGGTGAGCATCTGCGCGAGATCCG 443
DB 247 CCGGCGGGGTGTGCGCGGTGAGATCGAGCGGAGGTGAGCATCTGCGCGAGATCCG 306
QY 444 CCACCCCAATCATCATCGCTGACGATGTGTTCGAGACAACAACAGATGTGTCTGAT 503
DB 307 CCACCCCAATCATCATCGCTGACGATGTGTTCGAGACAACAACAGATGTGTCTGAT 366
QY 504 CTTGGA 509
DB 367 CTTGGA 372

RESULT 11
BG663786 411 bp mRNA linear EST 30-APR-2001
LOCUS BG663786
DEFINITION DRAAANF02 Rat DRG Library Rattus norvegicus cDNA clone DRAAANF02
5', mRNA sequence.
ACCESSION BG663786
VERSION BG663786.1 GI:13885708
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 411)
AUTHORS Xiao,H.S., Huang,Q.H., Zhang,F.X., Bao,L., Lu,Y.J., Guo,C.,
Yang,L., Huang,W.J., Fu,G., Xu,S.H., Cheng,X.P., Yan,Q., Zhu,Z.D.,
Zhang,X., Chen,Z., Han,Z.G. and Zhang X.
TITLE Identification of gene expression profile of dorsal root ganglion
in the rat peripheral axotomy model of neuropathic pain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (12), 8360-8366 (2002)
MEDLINE 22056133
PUBMED 12060780
COMMENT Contact: Zhang Xu

Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P.R. China
Tel: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn
This clone is also available at Chinese National Human Genome
Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
Pudong New Area, P.R. China. Please contact with Zhang Xu
(xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@chgc.sh.cn)
PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLYA-No.
FEATURES
source Location/Qualifiers
1..411
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRAAANF02"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
/clone_id="Rat DRG Library"
ORIGIN
Query Match 11.3%; Score 123; DB 12; Length 411;
Best Local Similarity 99.1%; Pred. No. 2.2e-45;
Matches 223; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 556 CACTGACAGAGATGAGCGCCAGCAGTTCCTCAAGACAGATCTGACGGTGTCCACTACC 615
DB 1 CACTGACAGAGATGAGCGCCAGCAGTTCCTCAAGACAGATCTGACGGTGTCCACTACC 60
QY 616 TGCACCTCAAGCGCATCCGCACTTTGACCTGGAAGCCGGGAACATATGTGTGACACA 675
DB 61 TGCACCTCAAGCGCATCCGCACTTTGACCTGGAAGCCGGGAACATATGTGTGACACA 120
QY 676 AGCATGACAGCCGCCACGATTAAGTCATCATGCTTTGGCATCGCGACAGATCGAGG 735
DB 121 AACATGACAGCCGCCACGATTAAGTCATCATGCTTTGGCATCGCGACAGATCGAGG 180
QY 736 CCGGTAGCGAGTTCAGAAACATCTTTGGACAGCCAGAGTTGTCTG 780
DB 181 CCGGTAGCGAGTTCAGAAACATCTTTGGACAGCCAGAGTTGTCTG 225

RESULT 12
CB787038 451 bp mRNA linear EST 16-MAY-2003
LOCUS CB787038
DEFINITION AMGNNUC:MRBE3-00123-E9-A rat brain E15 (10374) Rattus norvegicus
cDNA clone mrep3-00123-e9 5', mRNA sequence.
ACCESSION CB787038
VERSION CB787038.1 GI:28875429
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00123 row: e column: 9.
FEATURES
source Location/Qualifiers
1..451

ORIGIN

Query Match 9.4%; Score 103; DB 14; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.8e-39;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mbed3-00123-e9"
/tissue_type="brain E15"
/clone_lib="rat brain E15 (10374)"
/note="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat brain E15"

QY 288 CAGCGGCGAGTGGCGCATGCGGCGAGTGCAGCAGAGGAGGACCGGCGATGAGTACGC 347
Db 196 CAGCGGCGAGTGGCGCATGCGGCGAGTGCAGCAGAGGAGGACCGGCGATGAGTACGC 255

QY 348 GGCCTAAGTCTATAAGAGCGCGCGCTGCCGTCAGCGCGCGC 390
Db 256 GGCCTAAGTCTATAAGAGCGCGCGCTGCCGTCAGCGCGCGC 298

RESULT 13 136 bp mRNA linear EST 19-MAY-1998
AA966990
LOCUS ua38n03.r1 Soares mammary gland NbMWG Mus musculus cDNA clone
DEFINITION IMAGE:1349045 5' similar to SW:DAPK_HUMAN P53355 DEATH-ASSOCIATED
PROTEIN KINASE 1; mRNA sequence.
AA966990
VERSION AA966990.1 GI:3140883
SOURCE Mus musculus (house mouse)
KEYWORDS EST.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 136)
AUTHORS Marra,M., Hillier,J., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:697837

FEATURES
source
1. 136
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1349045"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NbMWG"
/note="Organ: mammary gland; Vector: pRT3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5'
TGTTCACATCTGAGAGTGGAGCGCGCCCGAGATGTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRT3D vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Benito Soares and M. Fatima
Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 136;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 CCCATTAGCTCATGACTTTGGCATTCGCCGACAGATCGAGGC 736
Db 80 CCCATTAGCTCATGACTTTGGCATTCGCCGACAGATCGAGGC 123

RESULT 14 216 bp mRNA linear EST 21-JAN-1997
AA027503
LOCUS mi04ell.r1 Soares mouse placenta 4NBMP13.5 14.5 Mus musculus cDNA
DEFINITION Clone IMAGE:459500 5' similar to PIR:S39269 S39269 DAP-Kinase -
human; mRNA sequence.
AA027503
VERSION AA027503.1 GI:1493497
SOURCE EST.
ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 216)
AUTHORS Marra,M., Hillier,J., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:276388

FEATURES
source
1. 216
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:459500"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares mouse placenta 4NBMP13.5 14.5"
/note="Organ: placenta; Vector: pRT3D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'
TGTTCACATCTGAGAGTGGAGCGCGCCCGAGATGTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT3D vector. Library
was constructed through one round of normalization, and was

ORIGIN

constructed by Bento Soares and M.Fatima Bernaldo."

Query Match 4.0%; Score 44; DB 9; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 CGCATTAAGCTCATCGACTTGGCGCGCACAGATCGAGCG 736

Db 44 CGCATTAAGCTCATCGACTTGGCGCGCACAGATCGAGCG 87

RESULT 15

AA027581

LOCUS 235 bp mRNA linear EST 21-JAN-1997
DEFINITION m106c09.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
clone IMAGE:459664 5' similar to PIR:S39269 S39269 DAP-kinase -
human ; mRNA sequence.

AA027581.1 GI:1493592

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
Mus musculus (house mouse)

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 235)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterson, R.

TITLE
JOURNAL
COMMENT

The WashU-HMT Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LIND, contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:276552
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 205.

FEATURES

source

location/Qualifiers

1..235
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:459664"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares mouse placenta 4NDMP13.5 14.5"
/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTACCAATCTGAAGTGGAGCGCGCGCAATTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M.Fatima Bernaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 235;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGGAGATCCGCCACCCCAACATCAT 458
Db 80 CGCGAGGTGAGCATCTCGCGGAGATCCGCCACCCCAACATCAT 123

RESULT 16

AA864020

LOCUS 316 bp mRNA linear EST 11-MAR-1998
DEFINITION vx88a02.r1 Soares_thymus_2NDMT Mus musculus cDNA clone
IMAGE:1282250 5' similar to SW:BNPK_HUMAN P5355 DEATH-ASSOCIATED
PROTEIN KINASE 1 ; mRNA sequence.

AA864020.1 GI:2956499

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
Mus musculus (house mouse)

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 316)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterson, R.

TITLE
JOURNAL
COMMENT

The WashU-HMT Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LIND, contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:674050
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES

source

location/Qualifiers

1..316
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1282250"
/sex="male"
/tissue_type="thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares_thymus_2NDMT"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCGTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bernaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGGAGATCCGCCACCCCAACATCAT 458
Db 185 CGCGAGGTGAGCATCTCGCGGAGATCCGCCACCCCAACATCAT 228

[illegible]

VERSION	AA013814.1	GI:1474889
KEYWORDS	EST.	Mus musculus (house mouse)
SOURCE	Mus musculus	
ORGANISM	Mus musculus	
REFERENCE	1 (bases 1 to 340)	Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kudaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellander, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Teisberg, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
AUTHORS		The WashU-HMNI Mouse EST Project
TITLE	Unpublished (1996)	
JOURNAL		
COMMENT	Contact: Marras, M/Mouse EST Project WashU-HMNI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. MG:267057	
FEATURES	Seq primer: EPPRimer.	
SOURCE	Location/Qualifiers	
	1..340	/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="C57BL/6J"
		/db_xref="taxon:10090"
		/clone="IMAGE:441721"
		/sex="unknown"
		/tissue_type="placenta"
		/dev_stage="adult"
		/lab_host="DH10B"
		/clone_id="Soares mouse placenta 4NBMP13.5 14.5"
		/note="Organ: placenta; Vector: pTR73D-Pac (Pharmacla) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',
		TGTTCACATCTCATCGACTTGGCATGGCCACAGATCGAGC 736
		T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacla), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN		
	Query Match	4.0%; Score 44; DB 9; Length 340;
	Best Local Similarity	100.0%; P-Id. No. 2.6e-10;
	Matches 44; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	693 CGCATTAAGCTCATCGACTTGGCATGGCCACAGATCGAGC 736	
Db	48 CGCATTAAGCTCATCGACTTGGCATGGCCACAGATCGAGC 91	
RESULT 19		
LOCUS	BE654598	343 bp mRNA linear EST 06-SEP-2000
DEFINITION	UT-M-AJI-agr-e-10-0-UI. r1 NIH BMAP MOB N Mus musculus cDNA clone	
ACCESSION	BE654598	
VERSION	BE654598.1	GI:9980511
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryotes: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 343)	Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene	

JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PUBMED
8889548
COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: m8t@mail.nih.gov

CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP CDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
1..343
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-AJ1-ag2-e-10-0-UR"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP MOB N"
/note="Vector: pT7T3-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH BMAP MOB N library is a normalized library constructed
from mouse olfactory bulbs. The tag is a string of 5
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996. Tissue provided by Ms. Annie Novakovich,
Zivic-Miller Laboratories."

ORIGIN

Query Match 4.0%; Score 44; DB 10; Length 343;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCAACCACATCAT 458
DB 216 CGCGAGGTGAGCATCTCGCGAGATCCGCCAACCACATCAT 259

RESULT 20
AA982305 354 bp mRNA linear EST 27-MAY-1998
LOCUS UA53b01.r1 Soares_thymus_2NbWT Mus musculus cDNA clone
IMAGE:1361161 5' similar to SW:DAPK_HUMAN P53355 DEXTH-ASSOCIATED
PROTEIN KINASE 1 ; mRNA sequence.
AA982305
ACCESSION AA982305.1 GI:3160964
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 354)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisell,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The Mashu-HMT Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washu-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

FEATURES

source

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LINT; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:699201
Possible reversed clone: similarly on wrong strand
Seq primer: -28m3 rev2 ET from Amersham
High quality sequence stop: 232.
Location/Qualifiers
1..354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1361161"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares thymus 2NbWT"
/note="Vector: pT7T3-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5'
GTGACCATCTGAGTGGAGCGCGCGCTTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization and was constructed by Bento
Soares and W.Falma Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCAACCACATCAT 458
DB 301 CGCGAGGTGAGCATCTCGCGAGATCCGCCAACCACATCAT 344

RESULT 21
AA023436 376 bp mRNA linear EST 21-JAN-1997
LOCUS mh2f07.r1 Soares mouse placenta 4NDWP13.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:456517 5' similar to PIR:S39269 S39269 DAP-Kinase -
human ; mRNA sequence.
AA023436
ACCESSION AA023436.1 GI:1487369
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 376)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisell,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The Mashu-HMT Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washu-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LINT; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:273405

/lab host="DH10B"
 /clone.lib="Soares mouse embryo NbME13.5 14.5"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minori Ko, Wayne
 State Univ., from 2], double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo. "

ORIGIN

Query Match 4.0%; Score 44; DB 14; Length 395;
 Best Local Similarity 100.0%; Pred.No.2.7e-10;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 415 CCGGAGTGAGCATCTCGCGAGATCCGCCACCCCAATCAT 458
 |||||||
 Db 302 CCGGAGTGAGCATCTCGCGAGATCCGCCACCCCAATCAT 345

RESULT 24 397 bp mRNA linear EST 21-JAN-1997
 LOCUS AA020495
 DEFINITION mh1904.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 clone IMAGE:455478 5' similar to PIR:S39269 S39269 DAP-kinase -
 human ;, mRNA sequence.

ACCESSION AA020495
 VERSION AA020495.1 GI:1484285
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 397)

REFERENCE

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Stepce,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

FEATURES

Source

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:272366
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 355.

Location/Qualifiers

1..397
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:455478"
 /sex="unknown"
 /tissue_type="placenta"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone.lib="Soares mouse placenta 4NDMP13.5 14.5"
 /note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
 with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 1st strand cDNA was primed with a Not I - oligo(dT) primer

[5',
 TGTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M.Fatima Bonaldo. "

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 397;
 Best Local Similarity 100.0%; Pred.No.2.7e-10;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 415 CCGGAGTGAGCATCTCGCGAGATCCGCCACCCCAATCAT 458
 |||||||
 Db 220 CCGGAGTGAGCATCTCGCGAGATCCGCCACCCCAATCAT 263

RESULT 25 399 bp mRNA linear EST 21-JAN-1997
 LOCUS AA013793
 DEFINITION mh1904.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 clone IMAGE:442398 5' similar to PIR:S39269 S39269 DAP-kinase -
 human ;, mRNA sequence.

ACCESSION AA013793
 VERSION AA013793.1 GI:1474859
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 399)

REFERENCE

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Stepce,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:267734
 Seq primer: EMP1mer
 High quality sequence stop: 342.

Location/Qualifiers

1..399
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:442398"
 /sex="unknown"
 /tissue_type="placenta"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone.lib="Soares mouse placenta 4NDMP13.5 14.5"
 /note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
 with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 1st strand cDNA was primed with a Not I - oligo(dT) primer
 [5',
 TGTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M.Fatima Bonaldo. "

```

ORIGIN
Query Match          4.0%: Score 44; DB 9; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      693 CGCATTAAGCTCATGACCTTTGGATCCGCGACAGATCGAGGC 736
      |||
      108 CGCATTAAGCTCATGACCTTTGGATCCGCGACAGATCGAGGC 151

RESULT 26
AL363655          400 bp      mRNA      linear      EST 04-AUG-2000
LOCUS            AL363655
DEFINITION       AL363655 ICRFP 522 and 523 Mus musculus cDNA clone K0322848 5',
ACCESSION        AL363655
VERSION          AL363655.1
KEYWORDS         GI:9692891
SOURCE           EST.
ORGANISM         Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bickhoff, L.H., Schuchard, J., Ivanov, I., Meier-Ewert, S., O'Brien, J.,
Malik, A., Tandon, N., Wolski, E., Rohls, E., Nyarsik, L.,
Reinhardt, R., Metfeld, W. and Lehrach, H.
Tissue gene expression analysis using arrayed normalized cDNA
libraries
Genome Res. (2000) In press
Contact: MPIMG
Abt. Lehrach
Max Planck Institut fuer Molekulare Genetik
Institute 73, Berlin, 14195 Germany
The cDNA libraries ICRFP 522 and 523 were normalized with
oligonucleotide fingerprinting, resulting in a unique subset of
5376 cDNA clones.
Location/Qualifiers
1..400
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="K9322848"
/tissue_type="embryo"
/dev_stage="9 and 12 pc embryo"
/clone_id="ICRFP 522 and 523"

FEATURES
Source
1..400
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="K9322848"
/tissue_type="embryo"
/dev_stage="9 and 12 pc embryo"
/clone_id="ICRFP 522 and 523"

ORIGIN
Query Match          4.0%: Score 44; DB 9; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      415 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
      |||
      252 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 295

RESULT 27
BE853739          405 bp      mRNA      linear      EST 27-SEP-2000
LOCUS            BE853739
DEFINITION       BE853739 Soares thymus 2bMT Mus musculus cDNA clone
IMAGE:347039 5' similar to TR:054784 054784 DEATH-ASSOCIATED
KINASE 3 ;, mRNA sequence.
BE853739
VERSION          BE853739.1
KEYWORDS         GI:10312309
SOURCE           EST.
ORGANISM         Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 405)
NCI-CCAF ntcp://www.ncbi.nlm.nih.gov/ncicgap.
REFERENCE
1 (bases 1 to 405)
NCI-CCAF ntcp://www.ncbi.nlm.nih.gov/ncicgap.

```

TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
JOURNAL
 Unpublished (1997)
COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgsbbs-remail.nih.gov
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1313139
 Seq primer: -40RP from Glibco
 High quality sequence stop: 348.
FEATURES
 source
 1..405
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:3472039"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH103"
 /note="vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCAATCTGAAATGAGGAGCGCGGCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pRT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Benito
 Soares and M. Fatima Bernaldo."

Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@ssc.riken.go.jp,
URL: http://genome.ssc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matsuhira, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsushita, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.ssc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES

source

Location/Qualifiers
1..422
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F91012A07"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"

ORIGIN

Query Match 4.0%; Score 44; DB 10; Length 422;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
DB 308 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 351

RESULT 29
AA050632 423 bp mRNA linear EST 09-SEP-1996
LOCUS m18e05.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:476480 5' similar to PIR:S39269 S39269 DAP-kinase -
human ;, mRNA sequence.

ACCESSION
VERSION AA050632.1 GI:1530305
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 423)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HM1 Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:287224

Putative full length read
vector to vector length is 475
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 401.
Location/Qualifiers

1..423
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:476480"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/clone_lib="Soares mouse embryo NBME13.5 14.5"
/note="vector: p173D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGGGAATTTTCTTTTCTTTT
T 3']; on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko; Wayne
State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p173 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
DB 242 CGCGAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 285

RESULT 30
W85231 423 bp mRNA linear EST 12-SEP-1996
LOCUS m152b04.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:408655 5' similar to PIR:S39269 S39269 DAP-Kinase -
human ;, mRNA sequence.

ACCESSION
VERSION W85231.1 GI:1541591
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 423)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HM1 Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397701.
Contact: Marra M/Mouse EST Project

WashU-HM1 Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source

MGI:252423
Seq primer: -28M13 rev2 from Amer sham
High quality sequence stop: 410.
Location/Qualifiers

1..423
/organism="Mus musculus"
/mol_type="RNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:408655"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAGTGGAGCGCGCGGAATTTTCTTTTCTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
State Univ., from 2); double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pRT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 14; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

693 CGCATTAGCTCATCGACTTGGCATCGCCACAGATCGAGCG 736
DB 251 CGCATTAGCTCATCGACTTGGCATCGCCACAGATCGAGCG 294

RESULT 31
LOCUS BY040500 428 bp mRNA linear EST 06-DEC-2002
DEFINITION BY040500 RIKEN full-length enriched, pooled tissues.
cell_line=TRB-55B88, etc. Mus musculus cDNA clone I730015K16 5',
mRNA sequence.

ACCESSION BY040500
VERSION BY040500.1 GI:26145943
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 428)

Okazaki, Y., Futuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamakata, I.,
Kiyosawa, H., Isgl, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Balderelli, R., Hill, D.P., Bull, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Karapin, A., Matsuda, H.,
Batilov, S., Beisel, K.W., Blake, J.A., Brad, D., Bruscia, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragan, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gasteier, T.,
Gerhold, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawai, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Konagaya, A.,
Kurochik, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pettes, G., Pesole, G.,
Petrovski, N., Pillai, R., Portius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringswald, M.,
Sandelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K.,
Sultana, R., Takekawa, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Vizarro, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

TITLE

JOURNAL MEDLINE
PUBMED
COMMENT
Contract: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numasaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Wataniki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia. Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES

source

1..428
/organism="Mus musculus"
/mol_type="RNA"
/db_xref="taxon:10090"
/clone="I730015K16"
/clone_lib="RIKEN full-length enriched, pooled tissues,
cell_line=TRB-55B88, etc."
/note="pooled tissues: (sex=mix, cell_line=TRB-55B88),
(sex=mix, cell_line=CRL-1722 L5178Y-R)"

ORIGIN

Query Match 4.0%; Score 44; DB 13; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

415 CCGGAGGTGAGATCTCTCGCGAGATCGCCACCCACATCAT 458
DB 295 CCGGAGGTGAGATCTCTCGCGAGATCGCCACCCACATCAT 338

RESULT 32
LOCUS BF147326 437 bp mRNA linear EST 26-OCT-2000
DEFINITION BF147326 uweh08.y1 Soares mammary gland NMLWS Mus musculus cDNA clone
IMAGE:3467199 5' similar to TR:054784 054784 DEATH-ASSOCIATED
KINASE 3, mRNA sequence.
ACCESSION BF147326
VERSION BF147326.1 GI:11028721

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 437)
AUTHORS NC-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: uw68h08.xl
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1386559
Seq primer: -40RP from Gldco
High quality sequence stop: 402.
Location/Qualifiers
1..437
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3467139"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - Gldco(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 4.0%; Score 44; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCGCGACCCCAACATCAT 458
|||||
DB 279 CGCGAGGTGAGCATCTCGCGAGATCGCGACCCCAACATCAT 322

RESULT 33
W77571 439 bp mRNA linear EST 20-JUN-1996
LOCUS me767e2.r1 Soares mouse embryo NBM3.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:400654 5' similar to PIR:S39269 S39269 DAP-kinase -
human ; mRNA sequence.
W77571
W77571.1 GI:1387622
EST.
Mus musculus (house mouse)
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 439)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Stepcoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810
Email: mouseest@atson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:244422
Seq primer: EPPRimer
High quality sequence stop: 341.
Location/Qualifiers
1..439
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:400654"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/clone_lib="Soares mouse embryo NBM3.5 14.5"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 4.0%; Score 44; DB 14; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTCGCGAGATCGCGACCCCAACATCAT 458
|||||
DB 135 CGCGAGGTGAGCATCTCGCGAGATCGCGACCCCAACATCAT 178

RESULT 34
AA475990 443 bp mRNA linear EST 18-JUN-1997
LOCUS vH25912.r1 Soares mammary_gland NBM3.5 14.5 Mus musculus cDNA clone
DEFINITION IMAGE:876550 5' similar to TR:G434847 G434847 DAP-KINASE.; mRNA
sequence.
AA475990
AA475990.1 GI:2203841
EST.
Mus musculus (house mouse)
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 443)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Stepcoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@atson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:516030

Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 228.

Location/Qualifiers

FEATURES

SOURCE

```
1.443
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:876550"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NbmMG"
/notes="Organ: mammary gland; Vector: pRTT3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTACCAATCTGAGATGAGGCGCGCGGAGATGTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRTT3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
```

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 443;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCGAGGTGAGCATCTGCGGAGATCGCCACCCCAATCAT 458
|||||
Db 249 CGCGAGGTGAGCATCTGCGGAGATCGCCACCCCAATCAT 292

RESULT 35
AA023330 446 bp mRNA linear EST 21-JAN-1997
LOCUS m71b08.r1 Soares mouse placenta 4NBMP13.5 14.5 Mus musculus CDNA
DEFINITION clone IMAGE:456375 5' similar to PIR:S39269 S39269 DAP-kinase -
human; mRNA sequence.

ACCESSION AA023330
VERSION AA023330.1 GI:1487263
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 446)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:273263
Seq primer: -28m13 rev2 from Amersham
High quality sequence stop: 406.

FEATURES

SOURCE

```
1.446
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:876550"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
```

```
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:456375"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares mouse placenta 4NBMP13.5 14.5"
/notes="Organ: placenta; Vector: pRTT3D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'
TGTACCAATCTGAGATGAGGCGCGCGGAGATGTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRTT3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 446;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 CGCATTAAGCTCATGCACTTTGGATCGCGGACAGATGAGGC 736
|||||
Db 91 CGCATTAAGCTCATGCACTTTGGATCGCGGACAGATGAGGC 134

RESULT 36
A1180838 446 bp mRNA linear EST 08-OCT-1998
LOCUS ub86b04.r1 Soares thymus_2NBMT Mus musculus CDNA clone
DEFINITION IMAGE:1395343 5' similar to SW:DAFK_HUMAN P53355 DEATH-ASSOCIATED
PROTEIN KINASE 1; mRNA sequence.

ACCESSION A1180838
VERSION A1180838.1 GI:3731476
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 446)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:907059
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 432.

FEATURES

SOURCE

```
1.446
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1395343"
/sex="male"
/tissue_type="thymus"
/dev_stage="4 weeks"
```

/lab_host="DH10B"
/clone_id="Soares thymus 2NbMT"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5].
TGTTACCAATCTGAGTGGAGCGCGCGCTTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 446;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
Db 239 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 282

RESULT 37

BF012286 449 bp mRNA linear EST 06-OCT-2000
LOCUS U54401.1 Soares_NKMD_mandible Mus musculus cDNA clone
IMAGE:3514056 5' similar to TR:043293 O43293 ZIP-KINASE. ; mRNA
sequence.

ACCESSION BF012286
VERSION BF012286
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 449)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicap.

REFERENCE

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL

Other_ESTs: u54401.x1
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov

COMMENT

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:1396936
Seq primer: -40RP from Gibco
High quality sequence stop: 441.
Location/Qualifiers

FEATURES

source

1..449
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_image="3514056"
/tissue_type="mandible"
/lab_host="DH10B (phage-resistant)"
/clone_id="Soares NKMD mandible"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: NotI; Site 2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5].
TGTTACCAATCTGAGTGGAGCGCGCGCTTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 10; Length 449;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 458
Db 196 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAACATCAT 239

RESULT 38

AA023742 454 bp mRNA linear EST 21-JAN-1997
LOCUS mh81b03.r1 Soares mouse placenta 4NMPI3.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:457325 5' similar to TR:0434847 G434847 DAP-KINASE. ;
RNA sequence.
AA023742
ACCESSION AA023742.1 GI:1487657
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 454)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucada, T., Lacy, M., Le, M., Martin, U., Morris, M.,
Schellander, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu

JOURNAL

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:274213
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 450.
Location/Qualifiers

FEATURES

source

1..454
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="457325"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="Soares mouse placenta 4NMPI3.5 14.5"
/note="Organ: Placenta; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker. Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5].
TGTTACCAATCTGAGTGGAGCGCGCGCGAATTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 454;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

693 CGCATTAAGCTCATGACTTGGATCGCGAGATCGAGATCGAGGC 736
Db 109 CGCATTAAGCTCATGACTTGGATCGCGAGATCGAGATCGAGGC 152

RESULT 39

AA017794

LOCUS AA017794 456 bp mRNA linear EST 21-JAN-1997
 DEFINITION m447909.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 clone IMAGE:445696 5' similar to PIR:S39269 S39269 DAP-kinase -
 human ; mRNA sequence.

ACCESSION AA017794
 VERSION AA017794.1 GI:1481180
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 456)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:271032
 Seg primer: -28M13 rev2 from Amersham
 High quality sequence stop: 455.
 Location/Qualifiers

FEATURES
 source
 1..456
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:445696"
 /sex="unknown"
 /tissue_type="placenta"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Soares mouse placenta 4NDMP13.5 14.5"
 /note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
 with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 1st strand cDNA was primed with a Not I - oligo(dT) primer
 [5'
 TGTACCAATCTGAGTGGAGCGCGCGGAAATTTTCTTTTCTTTTCTTTTCTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 CGCATTAAGCTCATCGACTTGGCATCGCGACAGATCGAGC 736
 Db 108 CGCATTAAGCTCATCGACTTGGCATCGCGACAGATCGAGC 151

RESULT 40
 LOCUS AA015164
 DEFINITION m457806.r1 Soares mouse embryo NDME13.5 14.5 Mus musculus cDNA
 clone IMAGE:468226 5' similar to PIR:S39269 S39269 DAP-kinase -
 human ; mRNA sequence.

ACCESSION AA015164
 VERSION AA015164.1 GI:1476195
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 457)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:282042
 Seg primer: -28M13 rev2 from Amersham
 Possible reversed clone; similarity on wrong strand
 High quality sequence stop: 448.
 Location/Qualifiers

FEATURES
 source
 1..457
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:468226"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /clone_lib="Soares mouse embryo NDME13.5 14.5"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAGTGGAGCGCGCGGAAATTTTCTTTTCTTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Mincru Ko, Wayne
 State Univ., from 2] ; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 457;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAATCAT 458
 Db 293 CGCAGGTGAGCATCTCGCGAGATCCGCCACCCCAATCAT 336

RESULT 41
 LOCUS AA016932
 DEFINITION m459808.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 clone IMAGE:444724 5' similar to PIR:S39269 S39269 DAP-kinase -
 human ; mRNA sequence.

ACCESSION AA016932
 VERSION AA016932.1 GI:1478564
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:290272
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 442.
 Location/Qualifiers
 1..475
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:479528"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /clone_lib="Soares mouse embryo NBM13.5-14.5"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGCGCCGAGATTTTCTTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos (total RNA provided by Marnett K., Wayne
 State Univ., from 2 1); double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 p773 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M. Fatima Bonaldo. "

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-10; Indels 0; Gaps 0;
 Matches 44; Conservative 0; Mismatches 0;

QY 693 CGCATTAGCTCATGACTTTGGCATCGGCACAGATCGAGGC 736
 127 CGCATTAGCTCATGACTTTGGCATCGGCACAGATCGAGGC 170

RESULT 44 482 bp mRNA linear EST 25-AUG-2000
 BE629354
 LOCUS
 DEFINITION
 u337a01.y1 Soares mammary gland NMLMG Mus musculus cDNA clone
 IMAGE:3374088 5' similar to TR:054784 054784 DEATH-ASSOCIATED
 KINASE 3 ; mRNA sequence.
 BE629354
 ACCESSION BE629354.1 GI:9912042
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 (bases 1 to 482)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1083692
 Seq primer: -40RP from Gibco
 High quality sequence stop: 435.
 Location/Qualifiers

FEATURES

source
 1..482
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:3374088"

ORIGIN

Query Match 4.0%; Score 44; DB 10; Length 482;
 Best Local Similarity 100.0%; Pred. No. 2.9e-10; Indels 0; Gaps 0;
 Matches 44; Conservative 0; Mismatches 0;

QY 693 CGCATTAGCTCATGACTTTGGCATCGGCACAGATCGAGGC 736
 42 CGCATTAGCTCATGACTTTGGCATCGGCACAGATCGAGGC 85

RESULT 45 483 bp mRNA linear EST 11-JUN-1997
 AA467508
 LOCUS
 DEFINITION
 ve01c12.r1 Soares mouse NBMH Mus musculus cDNA clone IMAGE:808822
 5' similar to TR:G434847 G434847 DAF-KINASE ; mRNA sequence.
 AA467508
 ACCESSION AA467508.1 GI:2193648
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 (bases 1 to 483)
 Gaisel,S., Kucaba,T., Lacy,M., Le,M., Martin,U., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterson,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:485166
 Seq primer: -28M13 rev2 ET from Amersham
 High quality sequence stop: 472.
 Location/Qualifiers

FEATURES

source

/sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"
 /clone_lib="Soares mammary gland NMLMG"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified p773 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Bonaldo. "

1..483
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:808822"
 /sex="male"
 /tissue_type="heart"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares mouse NBMH"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGCGCCGAGATTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

ORIGIN

Query Match 4.0%; Score 44; DB 9; Length 483;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 693 CGCATTAGCTCATCGACTTTGGCAGCGCAGATCGAGGC 736
Db 76 CGCATTAGCTCATCGACTTTGGCAGCGCAGATCGAGGC 119

Search completed: June 30, 2004, 21:16:00
Job time : 3359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 01:24:35 ; Search time 522 Seconds
(without alignments)
3179.428 Million cell updates/sec

Title: FIGURE9
Perfect score: 1826
Sequence: 1 XMAIGLGRSPXGQDLS.....GXMSIGLIDMPCGPHQXRX 361

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3031105 seqs, 2298700234 residues
Total number of hits satisfying chosen parameters: 6062210

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xjh
-Q/cgnt2/USFTO.spool/GBRREXUS641/rnatc_29062004_143151_24597/app_query.fasta_1.519
-DB=Published.Applications_NA -Qfmt=fastap -SUFix=imdb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000
-USER=GBRREXUS641 @CGN 1.1 221 @rnatc_29062004_143151_24597 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESOURT -NRG_SCORES=0 -WAIT -DSPBLOCK=100 -JONCLOG
-DEV_TIMEOUT=120 -WANN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Published Applications_NA:*
- 2: /cgnt2_6/ptcodata/2/pubpna/US07_PUBCOMB.seq:*
- 3: /cgnt2_6/ptcodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgnt2_6/ptcodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgnt2_6/ptcodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgnt2_6/ptcodata/2/pubpna/PCUS_PUBCOMB.seq:*
- 7: /cgnt2_6/ptcodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgnt2_6/ptcodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgnt2_6/ptcodata/2/pubpna/US08_PUBCOMB.seq:*
- 10: /cgnt2_6/ptcodata/2/pubpna/US08_PUBCOMB.seq:*
- 11: /cgnt2_6/ptcodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgnt2_6/ptcodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgnt2_6/ptcodata/2/pubpna/US09_PUBCOMB.seq:*
- 14: /cgnt2_6/ptcodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgnt2_6/ptcodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgnt2_6/ptcodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgnt2_6/ptcodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgnt2_6/ptcodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgnt2_6/ptcodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	1746.5	95.6	1093	13	US-10-083-641A-6	Sequence 6, Appl1
2	632	2132	16	2132	US-10-262-445-118	Sequence 118, App
3	631.5	34.6	2224	15	US-10-059-585-11	Sequence 11, Appl
4	469.5	25.7	1716	13	US-10-087-192-1835	Sequence 1835, Ap
5	432.5	23.7	5910	13	US-10-342-887-1105	Sequence 1105, Ap
6	432.5	23.7	5910	13	US-10-045-400C-4	Sequence 4, Appl1
7	432.5	23.7	5910	13	US-10-172-118-1105	Sequence 1105, Ap
8	432.5	23.7	5910	16	US-10-133-937-12	Sequence 12, Appl
9	432.5	23.7	5910	16	US-10-352-684A-37	Sequence 37, Appl
10	432.5	23.7	5910	16	US-10-295-027-1018	Sequence 1018, Ap
11	432.5	23.7	5910	16	US-10-159-563-12	Sequence 12, Appl
12	432.5	23.7	5910	17	US-10-303-588-4	Sequence 4, Appl1
13	432.5	23.7	6046	17	US-09-971-352-65	Sequence 65, Appl
14	432.5	23.7	6046	15	US-10-252-157-452	Sequence 452, App
15	418	22.9	1592	13	US-10-087-192-1832	Sequence 1832, Ap
16	387.5	21.2	6099	16	US-10-062-674-1427	Sequence 1427, Ap
17	335.5	18.4	493	10	US-09-918-995-38039	Sequence 38039, A
18	295	16.2	438	10	US-09-918-985-1256	Sequence 1256, Ap
19	293.5	16.1	152330	13	US-10-087-192-1834	Sequence 1834, Ap
20	292.5	16.0	126413	13	US-10-087-192-1831	Sequence 1831, Ap
21	291.5	16.0	535	15	US-10-029-386-9853	Sequence 9853, Ap
22	278.5	15.3	491	9	US-09-864-761-560	Sequence 560, App
23	264	14.5	259	9	US-09-864-761-17354	Sequence 17354, A
24	263	14.4	402	10	US-09-918-995-6425	Sequence 6425, Ap
25	242	13.3	2980	13	US-10-423-543-10	Sequence 42, Appl
26	242	13.3	2980	13	US-10-354-358-91	Sequence 91, Appl
27	237.5	13.0	2337	13	US-10-423-543-12	Sequence 12, Appl
28	237.5	13.0	2337	13	US-10-116-326-1	Sequence 1, Appl1
29	237.5	13.0	2897	17	US-10-311-034-43	Sequence 43, Appl
30	222.5	12.2	1984	13	US-10-425-114-19039	Sequence 19039, A
31	222.5	12.2	2062	13	US-10-425-114-34609	Sequence 34609, A
32	222.5	12.2	2112	15	US-10-116-326-5	Sequence 5, Appl1
33	222.5	12.2	2218	10	US-09-820-780-1	Sequence 1, Appl1
34	222	12.2	2061	14	US-10-096-960-1	Sequence 1, Appl1
35	222	12.2	2061	17	US-10-623-505-1	Sequence 1, Appl1
36	221.5	12.1	1880	17	US-10-311-034-48	Sequence 48, Appl
37	221.5	12.1	3705	14	US-10-044-090-227	Sequence 227, App
38	219.5	12.0	1388	12	US-10-152-319A-1918	Sequence 1918, Ap
39	219.5	12.0	3015	13	US-10-296-115-590	Sequence 590, App
40	218.5	12.0	1074	13	US-10-024-036B-3	Sequence 3, Appl1
41	218.5	12.0	1074	16	US-10-446-175-3	Sequence 3, Appl1
42	218.5	12.0	1158	16	US-10-446-175-1	Sequence 1, Appl1
43	218.5	12.0	1467	13	US-10-425-114-16427	Sequence 16427, A
44	218.5	12.0	1495	17	US-10-649-400-1	Sequence 1, Appl1
45	218.5	12.0	1578	9	US-09-835-788A-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-10-083-641A-6
Publication No. US20030017568A1
GENERAL INFORMATION:
APPLICANT: HAYSTEAD, TIMOTHY A
TITLE OF INVENTION: SMOOTH MUSCLE MYOSIN PHOSPHATASE ASSOCIATED KINASE
FILE REFERENCE: 1579-647
CURRENT APPLICATION NUMBER: US/10/083, 641A
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/271,436
PRIOR FILING DATE: 2001-02-27
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 1093
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Putative nucleotide
OTHER INFORMATION: Sequence of smooth muscle MYPT-Kinase

NAME/KEY: Unsure
 LOCATION: (2), (37), (39), (1056), (1081) and (1092)
 OTHER INFORMATION: N can be A, C, G or T
 US-10-083-641A-6

Alignment Scores:

Pred. No.:	1,86e-162	Length:	1093
Score:	1746.50	Matches:	349
Percent Similarity:	96.41%	Conservative:	0
Best Local Similarity:	96.41%	Mismatches:	10
Query Match:	95.65%	Indels:	3
	13	Gaps:	2

FIGURE9 (1-361) X US-10-083-641A-6 (1-1093)

```

QY 2 Met***IleglyLeuIleglyArgSerSerPro**GlyGluLeuAspSerLeuSerAsp 21
DB 4 ATGATATCGGTTTAAATCGCGCGAGCTCGCCCNCGGCGAGCTGAGACTCCCTCTCAGAC 63
QY 22 LeuLeuLeuSerArgProGlnHisGlyLeuLeuLeuThr**LeuPheLeuGlyProArg 41
DB 64 CTCCTCTCTTCTCCCTCCAGCAGCGAGATTAACTCACTTGACTGTTCTTGCGTCCCGG 123
QY 42 CysArgAlaSerValLeuSerLeuLeuAlaIleProIysCysLeuSer**GlySerLeu 61
DB 124 TGCCGGCGCAGCGCTCTCTCTCCCTCAAGGCATCCCAAGTCTCTGATAGAGCTCTTG 183
QY 62 GlySerSerValValValGlyAsnLeuGlyThrAspAlaGlnIleGlyIleValVal 81
DB 184 GGCAGTTCTGTTGTGTGGGAAACCTGGGAACAGATCACAGAGCTGGGGGTACAGAGTC 243
QY 82 LeuProSerSerGlySerAlaAlaLeuSerCysSerPheProHisSerGly---PheAla 100
DB 244 CTGCTTCTCTGCGGTCTGCGACCGCTTACGTTCTCTTCCCTCCAGCGGCGAGTTGCGC 303
QY 101 IleValArgIysCys-----LysGlyThrGlyMetGlyIleValAlaIleValPheIleIys 118
DB 304 ATCGGCGCAGATCGCCAGCAGAGGCAAGGCAACCGGAGTGAAGTTCATGATTAAG 363
QY 119 LysArgArgLeuProSerSerArgArgIleValSerArgGluGluIleGluArgGluVal 138
DB 364 AAGCGCGCGCTCGCTCCAGCGCGCGGTGTGAGCGGTGAGGATCGAGCGAGAGGTG 423
QY 139 SerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPheGluAsn 158
DB 424 AGCATCCCGCGAGATCGCCAGCCCAACATCATCAAGCTGCGACAGATGATGAGAAC 483
QY 159 LysThrAspValValIleIleLeuGluLeuValSerGlyGlyGluLeuPheAspPheLeu 178
DB 484 AAGACAGATGTGGTGTGATCTTGAGCTGTGTGCGCGCGGCACTTTGCACTTTCTG 543
QY 179 AlaGluLysAspHis**GlnArgMetArgProArgSerSerSerSerArgSerTyrThr 198
DB 544 GCTGGAAGCATCTGACAGAGATAGAGGCAAGCAAGTCTCTCAACAGATCTCTGAGC 603
QY 199 ValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgArgThrSer 218
DB 604 GTGTCCACTACTGATCCAGAGCGCATCGCGCACTTGACCTGAACCCGAGAACATCA 663
QY 219 CysCysTyrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAlaSerArg 238
DB 664 TGTGCTGGAAGCATGACAGCCAGCCAGCCATTAAGCTCATGACTTTGGATCGCGC 723
QY 724 ACAGAGATCGAGCGCGGTGAGATTCAAGACATCTTTGGACCCGAGAGTCTGCGGTG 763
QY 259 ArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyLeuThrCys 278
DB 784 AAGGCGAGGTGTGGCGACACCGATAGAGTAGATTGACGCGCTTGCGTGAAGCTGCG 843
QY 279 LeuAsnAspProValPheHisSerPro**AspCysIleLeu**ThrThrTyrLeuGly 298

```

```

DB 844 CTCACAAATCTGCTTCCACAGCCCTGAGATTGTAACTATGAAACCATTCGCTTGA 903
QY 299 Ser***TyrValGluHisArgArgHisHisLeuHisProValSerAla**AspGlyGln 318
DB 904 AGCTGATATGTGAGCATTCGGGTGATCATCCATCTGAGAGTGCCTGAGATGGGAG 963
QY 319 GlyProGlnThrValProAlaArgGlyProGlyIleArgAlaGlyThrSerAlaAspCys 338
DB 964 GGGCTCAGACGTGTACTGCTTAGAGCCCGAGGATCAGGGCTGGGACCTCTGCAAACTGC 1023
QY 339 LysHisTyrGly**GluMetSerLeuGlyThrLeuAspMetProGlyProHisGln** 358
DB 1024 AAACACTGGGCGTGAAGATGTCTCTGGAAACNTGATATGCTGGGCGCCCAAGNT 1083
QY 359 ArgThr 360
DB 1084 AGCAGC 1089
RESULT 2
US-10-262-445-118
Sequence 118, Application US/10262445
Publication No. US20040014058A1
GENERAL INFORMATION:
APPLICANT: Alsodbrook II, John
APPLICANT: Burgess, Catherine
APPLICANT: Caterton, Elina
APPLICANT: Chant, John
APPLICANT: Chaudhuri, Amitabha
APPLICANT: Edinger, Solomon
APPLICANT: Gerlach, Valerie
APPLICANT: Glot, Loic
APPLICANT: Gorman, Linda
APPLICANT: Guo, Xiaojia
APPLICANT: Kekuda, Ramesh
APPLICANT: Mezes, Peter
APPLICANT: Millet, Isabelle
APPLICANT: Ooi, Chean Eng
APPLICANT: Patursajan, Meera
APPLICANT: Rieger, Daniel
APPLICANT: Spylek, Kimberly
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Zehnusen, Bryan
APPLICANT: Zhong, Haihong
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
FILE REFERENCE: 21402-462D
CURRENT APPLICATION NUMBER: US/10/262,445
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: 60/327,454
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,025
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,849
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/329,414
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 60/330,142
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 60/341,058
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/343,629
PRIOR FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 60/349,575
PRIOR FILING DATE: 2001-10-29
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 133
SOFTWARE: CuraSeqLast version 0.1
SEQ ID NO 118

```

```

; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)..(1455)
; US-10-262-445-118

Alignment Scores:
Pred. No.: 2,476-52 Length: 2132
Score: 632.00 Matches: 155
Percent Similarity: 62.69% Conservative: 13
Best Local Similarity: 57.84% Mismatches: 46
Query Match: 34.61% Indels: 54
Gaps: 9

FIGURE9 (1-361) x US-10-262-445-118 (1-2132)

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 163 TTTCGATCGTGGAGAGTCCCGCAGAGAGCGCGGCAAGAGTACGACCAAGTTC 222
QY 117 IleLysLysArgArgLeuProSerSerArgGlyValSerArgLysGluIleGluArg 136
Db 223 ATCAAGAGAGCGCGCGCTGATCCAGCGCGCGTGGGTGAGCCGGAGAGATCGAGCGG 282
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 283 GAGGTGAACATCCCTCGCGGAGATCCGGCACCCCAACATCATCACTCGACGACATCTTC 342
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 343 GAGAACAGACGACGCTGCTCTCATCTGAGCTGCTCTCGCGGAGCTCTTTGAC 402
QY 177 PheLeuIleGluLysAspHis***GlnArgMetArgProArgSerSerSerSerArg 196
Db 403 TTCTCGCGGAGAAAGAGTGTGCTGACGAGAGAGAGAGCCACAGTCTCTCAAGCAGATC 462
QY 196 rTPRThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
Db 463 CTGGACCGCGCTTCACTACCTGCACTTAAAGCGATCCACACTTTCAGACCGCGAA 522
QY 216 GlnSerCysCysThrProSerMetGlnProAlaHisAlaLeuSerSerSerThrLeu 236
Db 523 AACATCATGCTGCTGAGCAGAGAGCTCCACACCAAGATCAAGCTCATGCTTCGGG 582
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgLysSer 256
Db 583 ATGCGGACAGAGATCGAGCGGAGAGAGAGTTCAGAGACA-----TC 624
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGly 276
Db 625 TTTCGGACACCCCGAGATTGTGGC----- 647
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThr 296
Db 648 -----CCAGAGATTGTGAACATTCAGACCGCTG 675
QY 296 rPheLysLysSer***TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
Db 676 GGCTGTGAGCGGAGCATGTGAGCATGTGCTCATCACTATATCTT----- 722
QY 316 spgLyngLysProGln-----ThrValProAlaArgGlyProGlyIleArgAla 333
Db 723 -----CCTAGCGGTGATCCCGCTTCTCGGCGAGAC-----CAAGCAG 762
QY 333 lYthrSerAlaAsnCysLysHisHisTrpGly**GluMetSerLeu-----GlyThrLeu 351
Db 763 GAGAGCTGCAC-----CAACATCTCAGCGGTGAATCAAGATTTCAGAGAGATCTTC 816
QY 351 spMetProGlyProHisGln 357
Db 817 AGCAA-----CACCAAG 827

```

```

RESULT 3
US-10-059-585-11
; Sequence 11, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Furuhashi, Shin-ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(1576)
; US-10-059-585-11

Alignment Scores:
Pred. No.: 2,926-52 Length: 2224
Score: 631.50 Matches: 142
Percent Similarity: 69.12% Conservative: 8
Best Local Similarity: 65.44% Mismatches: 34
Query Match: 34.58% Indels: 33
Gaps: 3

FIGURE9 (1-361) x US-10-059-585-11 (1-2224)

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 284 TTTCGATCGTGGAGAGTCCCGCAGAGAGCGCGGCAAGAGTACGACCAAGTTC 343
QY 117 IleLysLysArgArgLeuProSerSerArgGlyValSerArgLysGluIleGluArg 136
Db 344 ATCAAGAGAGCGCGCGCTGATCCAGCGCGGTGGGTGAGCCGGAGAGATCGAGCGG 403
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 404 GAGGTGAACATCCCTCGCGGAGATCCGGCACCCCAACATCATCACTTCGACGACATCTTC 463
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 464 GAGAACAGACGAGCTGTCTCATCTGAGCTGTGCTGAGCGGAGAGCTCTTTGAC 523

```

```

QY 177 PheLeuAlaGluYsAspHis***GlnArgMetArgProArgSerSerSerArgSe 196
Db 524 TTCCTGGCGGAGAGAGATCGCTGACGAGACGAGCCACCCAGTTCCTCAAGCAGATC 583
QY 196 rTrpThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 584 CTGACGCGGCTTCACCTGACCTGCTTAAGCGCATGCGACACTTTAGCTGAAAGCCGGA 643
QY 216 gThrSerCysSTrPrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 644 AACATCATGCTGCTGACAGAAAGACGCTCCCAACCCAGATCAAGCTCATGACTTCGCG 703
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInSerSe 256
Db 704 ATCCGCGACAAAGATCGAGCGCGGGAACGAGTTCAAGACA-----TC 745
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
Db 746 TTCGGACACCCCGAGTTTGTCG----- 768
QY 276 eutHrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrT 296
Db 769 -----CCAGAGATGTGAACATGTGAGCCGCTG 796
QY 296 rPleuGlySer***TyrValGluHisArgArgHisHisLeuHisPro 311
Db 797 GGCCTGAGGCGGACATGTGAGCATGCTGCTCATCACTATATCCT 843

```

RESULT 4

```

US-10-087-192-1835
; Sequence 1835, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR APPLICATION NUMBER: US 09/799,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1835
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1835

```

```

Alignment Scores:
Pred. No.: 1.9e-36 Length: 1716
Score: 469.50 Matches: 112
Percent Similarity: 61.01% Conservative: 21
Best Local Similarity: 51.38% Mismatches: 52
Query Match: 25.71% Indels: 33
DB: 13 Gaps: 3

```

FIGURE9 (1-361) x US-10-087-192-1835 (1-1716)

```

QY 99 PheAlaAlaValAlaGlyLysCys-----LysGlyThrGlyMetGlyTyrAlaAlaLysPhe 116
Db 106 TTTGGCATCTGAGAGAAAGTGGCGGAGAAAGACGCGGCTTGAGTGTGACGCCAAGTTC 165
QY 117 IleuYsLysArgArgLeuProSerSerArgArgGlyAlaSerArgGluGluIleGluArg 136
Db 166 ATCAAGAACCGGACAGACCGGCGGCGGCTGTGAGCCCGGAGAGATGACGCG 225
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156

```

```

Db 226 GAGGTGACATCTCTGGCGGAGGTGCTGCCACCAATGTCATCAGCTGACGAGCTCTAT 285
QY 157 GluAsnLysTrpAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 286 GAGAACCCGACCGACGTGTGCTCTCATCTTGAAGTACGTGTGAGAGAGACTCTTCGAT 345
QY 177 PheLeuAlaGluYsAsp-His***GlnArgMetArgProArgSerSerSerArgSe 196
Db 346 TTCCTGGCGGAGAGAGATCTGAGTGAAGAGAGGCGCACACACTTCATTAAACAGATC 405
QY 196 rTrpThValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 406 CTGAGTGGGCTGAACTACTTTCACACAAAGAAATTCCTGCTCATTTGATTCACAGCCGAA 465
QY 216 gThrSerCysSTrPrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 466 AACATTAATGTTTGAAGAAGATTTCCATTCACACATCAAGCTGATGACTTTGGT 525
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInSerSe 256
Db 526 CTGCTCAGCAAAATGAGATGAGAGATTGAATTTAAG----- 562
QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
Db 563 -----ATATTTTGGAGACCCGGA 582
QY 276 eutHrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrT 296
Db 583 TTTGTGCTCCAGAAAT-----TGGAAGTACAGACCCCTG 618
QY 296 rPleuGlySer***TyrValGluHisArgArgHisHisLeuHisProVal 312
Db 619 GGTCTGAGCGCTGACATGTGAGCATGAGCGGTATCATCTCATCTCTCT 668

```

RESULT 5

```

US-10-342-887-1105
; Sequence 1105, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van t Veer, Laura Johanna
; APPLICANT: Van t Veer, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO: 1105
; LENGTH: 5910
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1105

```

```

Alignment Scores:
Pred. No.: 3.93e-32 Length: 5910
Score: 432.50 Matches: 106
Percent Similarity: 60.65% Conservative: 25
Best Local Similarity: 49.07% Mismatches: 54
Query Match: 23.69% Indels: 33
DB: 13 Gaps: 2

```

FIGURE9 (1-361) x US-10-342-887-1105 (1-5910)

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
 Db 406 TTTCGGGTGTGAAGAAATGCGGTGAGAAAGTACCGCCCTCCAGTATGCGCCCAAAATTC 465
 QY 117 IleLysLysArgLysLeuProSerSerArgArgLysValSerArgLysGluLysArg 136
 Db 466 ATCAAGAAAGAGAGACTAAGTCCAGCCGGGGGTGTGACCCCGAGACATCGACCGG 525
 QY 137 GluValSerLleLeuArgGluLleArgHisProAsnLleLleThrLeuHisAspValPhe 156
 Db 526 GAGTCAGCATCTCGAAGAGATCAGACCCCAATGTCATCACCCTGACGAGGTCTAT 585
 QY 157 GluAsnLysThrAspValIleLeuLleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
 Db 586 GAGAACAGACGAGCATCTGATCTTGAATCTGAGGTGCGAGCTGTTGAC 645
 QY 177 PheLeuAlaGluLysAspHis**Gln-ArgMetArgProArgSerSerSerArgse 196
 Db 646 TTCTTAGCTGAAAGAGATCTTTAAGTGAAGAGAGAACTGAATTTCTCAAAATTT 705
 QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
 Db 706 CTTAATGATGTTTACTACCTGACCTCCCTCAAAATCCCACTTGATCTTAAGCTGAG 765
 QY 216 GThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
 Db 766 AACATATGCTTTGGATAGAAATGTCGCCAACTCGAGATCAAGATCATTTGCTGGG 825
 QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgLysSerse 256
 Db 826 T----- 826
 QY 256 rSerValArgLysArgCysGlyHisHisProLleGly**IleLeuHisGlyLeuGlyLe 276
 Db 827 -----TGGCCCATTAATTTGACTTTGGAATGAATTAA 861
 QY 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrTr 296
 Db 862 AACATATTTGGG-ACTCAGAGTTTGTGCTCTCGA-GATAGTCAACTATGAACCTCTTG 919
 QY 296 PleuGlySer**TyrValGluHisArgArgHisHisLeuHisPro 311
 Db 920 GTCTTGAGGACAGATATGTGAGATCGGGGTATTAACCTATATCT 965
 RESULT 6
 US-10-045-400C-4
 ; Sequence 4, Application US/10045400C
 ; Publication No. US20030224509A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moon, Chulso
 ; APPLICANT: Mao, Li
 ; TITLE OF INVENTION: DAP-Kinase and HOXA9, Two Human Genes Associated with Genesis,
 ; FILE REFERENCE: 10620-101
 ; CURRENT APPLICATION NUMBER: US/10/045,400C
 ; CURRENT FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/250,083
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 5910
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (337)..(4632)
 ; OTHER INFORMATION:
 ; US-10-045-400C-4
 Alignment Scores: 3.93e-32 Length: 5910
 Pred. No.:

Score: 432.50 Matches: 106
 Percent Similarity: 60.65% Conservative: 25
 Best Local Similarity: 49.07% Mismatches: 54
 Query Match: 23.69% Indels: 33
 Gaps: 2
 DB: 13
 FIGURE9 (1-361) x US-10-045-400C-4 (1-5910)
 QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
 Db 406 TTTCGGGTGTGAAGAAATGCGGTGAGAAAGTACCGCCCTCCAGTATGCGCCCAAAATTC 465
 QY 117 IleLysLysArgLysLeuProSerSerArgArgLysValSerArgLysGluLysArg 136
 Db 466 ATCAAGAAAGAGAGACTAAGTCCAGCCGGGGGTGTGACCCCGAGACATCGACCGG 525
 QY 137 GluValSerLleLeuArgGluLleArgHisProAsnLleLleThrLeuHisAspValPhe 156
 Db 526 GAGTCAGCATCTCGAAGAGATCAGACCCCAATGTCATCACCCTGACGAGGTCTAT 585
 QY 157 GluAsnLysThrAspValIleLeuLleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
 Db 586 GAGAACAGACGAGCATCTGATCTTGAATCTGAGGTGCGAGCTGTTGAC 645
 QY 177 PheLeuAlaGluLysAspHis**Gln-ArgMetArgProArgSerSerSerArgse 196
 Db 646 TTCTTAGCTGAAAGAGATCTTTAAGTGAAGAGAGAACTGAATTTCTCAAAATTT 705
 QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
 Db 706 CTTAATGATGTTTACTACCTGACCTCCCTCAAAATCCCACTTGATCTTAAGCTGAG 765
 QY 216 GThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
 Db 766 AACATATGCTTTGGATAGAAATGTCGCCAACTCGAGATCAAGATCATTTGCTGGG 825
 QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgLysSerse 256
 Db 826 T----- 826
 QY 256 rSerValArgLysArgCysGlyHisHisProLleGly**IleLeuHisGlyLeuGlyLe 276
 Db 827 -----TGGCCCATTAATTTGACTTTGGAATGAATTAA 861
 QY 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrTr 296
 Db 862 AACATATTTGGG-ACTCAGAGTTTGTGCTCTCGA-GATAGTCAACTATGAACCTCTTG 919
 QY 296 PleuGlySer**TyrValGluHisArgArgHisHisLeuHisPro 311
 Db 920 GTCTTGAGGACAGATATGTGAGATCGGGGTATTAACCTATATCT 965
 RESULT 7
 US-10-172-118-1105
 ; Sequence 1105, Application US/10172118
 ; Publication No. US20030224374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yundong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999
 ; CURRENT APPLICATION NUMBER: US/10/172,118
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1105

```

; LENGTH: 5910
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_004938
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1105

Alignment Scores:
Pred. No.: 3,936-32 Length: 5910
Score: 432.50 Matches: 106
Percent Similarity: 60.65% Conservative: 25
Best Local Similarity: 49.07% Mismatches: 54
Query Match: 23.69% Indels: 33
Gaps: 2

FIGURE9 (1-361) x US-10-172-118-1105 (1-5910)

QY 99 PheAlaIleValArgLysCys-----LysGLYThrGlyMetGluTYrAlaAlaLysPhe 116
DB 406 TTTCGGGTGTGAAGAAATGCCGTGAGAAAGTACCGGCTCCAGTATGCCGCCAAATTTC 465
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 466 ATCAAGAAAAGAGAGACTAAGTCCAGCCGGGGGTGTGAGCCCGAGACATCGAGCGG 525
QY 137 GluValSerIleLeuArgGluIleArgHisProbsnIleIleThrLeuHisAspValPhe 156
DB 526 GAGGTCAAGCATCCGAAAGAGATCCAGACCCCAATGTCATCCACCTGCGAGAGTCTAT 585
QY 157 GluAsnLysThrAspValValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 586 GAGAACAAAGACGAGCGTATCTGATCTTGGAATCTCGTGAAGTGGCGAGCTGTTGAC 645
QY 177 PheLeuAlaGluLysAspHis***Gln-ArgMetArgProArgSerSerSerArgSe 196
DB 646 TTCTTAGCTGAAAAGAACTTTTAACGAAAGAGAACCACTGAATTCTCAACAAATT 705
QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 706 CTTAATGGTGTCTTACTACTGCTCACTCCCTTCAAAATCCGCCACTTGAATCTTAAGCCTGAG 765
QY 216 gThSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 766 AACATTAATGCTTTGGATAGAAATGTCGCCAACCTCGAGTCAAGATCATTTGACTTTGGG 825
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInserSe 256
DB 826 T----- 826
QY 256 rSerValArgGlyArgCysGlyHisAspProIleGly***IleLeuHisGlyLeuGlyLe 276
DB 827 -----TGGCCCATTAATAATTGACTTTGGAAATGAATTTAA 861
QY 276 uThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrTr 296
DB 862 AACATATTGGG-ACGCCAGAGTTTGTGCTCCCTGA-GATAGTCAACTATGAACCTCTTG 919
QY 296 PleuGlySer***TYrValGluHisArgArgHisHisLeuHisPro 311
DB 920 GTCTTGAGGACAGATATGTGAGTATCGGCGTAATTAACCTATATCTT 965

RESULT 8
US-10-133-937-12
; Sequence 12, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringer, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND

```

```

; TITLE OF INVENTION: OTHER BIOLOGICAL STATES
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133.937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 12
; LENGTH: 5910
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-12

Alignment Scores:
Pred. No.: 3,936-32 Length: 5910
Score: 432.50 Matches: 106
Percent Similarity: 60.65% Conservative: 25
Best Local Similarity: 49.07% Mismatches: 54
Query Match: 23.69% Indels: 33
Gaps: 2

FIGURE9 (1-361) x US-10-133-937-12 (1-5910)

QY 99 PheAlaIleValArgLysCys-----LysGLYThrGlyMetGluTYrAlaAlaLysPhe 116
DB 406 TTTCGGGTGTGAAGAAATGCCGTGAGAAAGTACCGGCTCCAGTATGCCGCCAAATTTC 465
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 466 ATCAAGAAAAGAGAGACTAAGTCCAGCCGGGGGTGTGAGCCCGAGACATCGAGCGG 525
QY 137 GluValSerIleLeuArgGluIleArgHisProbsnIleIleThrLeuHisAspValPhe 156
DB 526 GAGGTCAAGCATCCGAAAGAGATCCAGACCCCAATGTCATCCACCTGCGAGAGTCTAT 585
QY 157 GluAsnLysThrAspValValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 586 GAGAACAAAGACGAGCGTATCTGATCTTGGAATCTCGTGAAGTGGCGAGCTGTTGAC 645
QY 177 PheLeuAlaGluLysAspHis***Gln-ArgMetArgProArgSerSerSerArgSe 196
DB 646 TTCTTAGCTGAAAAGAACTTTTAACGAAAGAGAACCACTGAATTCTCAACAAATT 705
QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 706 CTTAATGGTGTCTTACTACTGCTCACTCCCTTCAAAATCCGCCACTTGAATCTTAAGCCTGAG 765
QY 216 gThSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 766 AACATTAATGCTTTGGATAGAAATGTCGCCAACCTCGAGTCAAGATCATTTGACTTTGGG 825
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgInserSe 256
DB 826 T----- 826
QY 256 rSerValArgGlyArgCysGlyHisAspProIleGly***IleLeuHisGlyLeuGlyLe 276
DB 827 -----TGGCCCATTAATAATTGACTTTGGAAATGAATTTAA 861
QY 276 uThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrTr 296
DB 862 AACATATTGGG-ACGCCAGAGTTTGTGCTCCCTGA-GATAGTCAACTATGAACCTCTTG 919
QY 296 PleuGlySer***TYrValGluHisArgArgHisHisLeuHisPro 311
DB 920 GTCTTGAGGACAGATATGTGAGTATCGGCGTAATTAACCTATATCTT 965

RESULT 9
US-10-352-684A-37
; Sequence 37, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Carroll, Joseph M.

```

```

APPLICANT: Healy, Aileen
APPLICANT: Welch, Nadine S.
APPLICANT: Kelly, Louise M.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212.
TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
FILE REFERENCE: ME102-0191P1NOMIIV
CURRENT APPLICATION NUMBER: US/10/352,684A
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 60/354,333
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: US 60/360,258
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/364,476
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/375,626
PRIOR FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: US 60/386,494
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: US 60/390,965
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: US 60/392,480
PRIOR FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: US 60/394,128
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: US 60/399,783
PRIOR FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: US 60/403,221
PRIOR FILING DATE: 2002-08-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 37
LENGTH: 5910
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (337)...(4632)
US-10-352-684A-37

Alignment Scores:
Pctd. No.: 3,93e-32 Length: 5910
Score: 432.50 Matches: 106
Percent Similarity: 60.65% Conservative: 25
Best Local Similarity: 49.07% Mismatches: 54
Query Match: 23.69% Indels: 33
DB: 16 Gaps: 2

FIGURE9 (1-361) x US-10-352-684A-37 (1-5910)
QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaIleLysPhe 116
DB 406 TTTGGCGTTGTGAAGAAATGCGGTGAGAAAGATGCGGCTCCAGATATCCCCCAATTC 465
QY 117 IleuLyArgArgLeuProSerSerArgGlyValSerArgGluGluArg 136
DB 466 ATCAAAAGAGGAGGCTAGTCCAGCCGCGGGGTAGCGGAGAGACATCAAGCGG 525
QY 137 GluValSerIleuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 526 GAGGTAGAGATCTGTAAGAGATCCAGCACCCCAATGATATCACTGACAGAGTCTAT 585
QY 157 GluAsnLyThrAspValValLeuIleLeuGluLeuValSerGlyGlyIleuPheAsp 176
DB 586 GAGAAACAACGACGCTCATCTCGATCTTGAACTCGTTCAGAGCGAGCTGTTGAC 645
QY 177 PheLeuAlaGluLysAspHis**Gln-ArgMetArgProArgSerSerSerArgSe 196
DB 646 TTCTTAGTGAAGAAATCTTAACTGAAGAGAGCACTGAATTTCTCAAAACAATT 705
```

```

QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
DB 706 CTAAATGGTGTACTTACTGACCTCCCTTCAAAATGCGCCACTTTGATCTTAAGCTTAG 765
QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaSerSerSerThrLeuAl 236
DB 766 AACATATAGCTTTTGAATAGAAATGTCCCAAACTCGAGATCAAGATTAAGCTTTGGG 825
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 826 T----- 826
QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyLe 276
DB 827 -----TGGCCCATAAATATGACTTTGAAATGAATTTAA 861
QY 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrTr 296
DB 862 AACATATTTGGG-ACCTCAGAGTTTGTGCTGCTCGA-GATAGTCAACTATGAACCTCTG 919
QY 296 PleuGlySer**TyrValGluHisArgArgHisHisLeuHisPro 311
DB 920 GTCTGAGCAGATATGTGAGTATCGGGGTAAATACCTATATCT 965

RESULT 10
US-10-295-027-1018
Sequence 1018, Application US/10295027
Publication No. US2003023250A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/340,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1018
LENGTH: 5910
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-1018
```



```

; ORGANISM: H. sapiens
;
; FEATURE:
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (337) ... (4632)
;
US-10-303-588-4

```

Alignment Scores:	
Pred. No.:	3 93e-32
Score:	42.50
Percent Similarity:	60.65%
Best Local Similarity:	45.07%
Query Match:	25.69%
DB:	17
Gaps:	2
Length:	5910
Matches:	106
Conservative:	25
Mismatches:	54
Indels:	33

FIGURE9 (1-361) X US-10-303-588-4 (1-5910)

```

QY      99  PheAlaIleValArgIysCys-----LysGlyThrGlyMetGluTyrAlaAlaIlePhe 116
Db      406  TTTCGGGTGTGAAGAAATGCGGTGAGAAAGTACCGGCTTCAGTATGCGCCAAATTC 465
QY      117  IleLysLysArgThrGluLeuProSerSerArgArgIValSerArgGluGluIleGluArg 136
Db      466  ATCAAGAAAAAGGAGGACTAAGTCCAGCCGGCGGGGTGTGAGCCGGAGAGACATGAGCGG 525
QY      137  GluValSerIleLeuArgGluIleArgHisProAlaIleIleThrLeuHisAspValPhe 156
Db      526  GAGGTGAGATCTGTGAAGAGATCCAGACCCCAATGATCATCACCCTGCAGAGAGTGTAT 585
QY      157  GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyLeuPheAsp 176
Db      586  GAGAACAGACGGAGCTCATCTGTATCTTTGGAACTCGTGAGAGGGGAGAGCTGTTTGAC 645
QY      177  PheLeuAlaGluLysAspHis**GlnArgMetArgProArgSerSerSerArgSe 196
Db      646  TTCTTTCGCGAAAAGGAAATCTTTAACTGAAGGAGAGCACTGAATTTCTCAACAAATT 705
QY      196  rTrpThrValSerThrThrCysThrProSerAlaSerArgThrIleuThr**SerArg 216
Db      706  CTTAATGCGTTTACTACCTGCAGCTCCCTTCAATGCCGCCACTTTGATCTTAAGCTGAG 765
QY      216  gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrIleu 236
Db      766  AACATATATCTTTTGATAGAAATGTCCCAAACTCGCATCAAGATCATGTAATTGGG 825
QY      236  aSerArgThrGlySerArgProValAlaSerSerArgThrSerIleuAlaArgGlnSer 256
Db      826  T----- 826
QY      256  rSerValArgGlyArgCysGlyHisIleProIleGly**IleLeuHisGlyLeuGlyLe 276
Db      827  -----TGGCCCTTAATAATTAAGTCTTGGAAATGAAATTAA 861
QY      276  uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**PheTrpTr 296
Db      862  AACATATTGGC-ACTCGAGAGTTGTGGCTCTCTA-GATGTCGAACATAAGAACTCTTG 919
QY      296  rLeuGlySer**TyrValGluHisArgArgHisHisIleuHisPro 311
Db      920  GTCCTGAGCGACATATGTGCACTATCGGGATATACATACTATATCT 965

RESULT 13
US-09-971-392-65
/ Sequence 65; Application US/09971392
/ Publication NO. US20030134283A1
/ GENERAL INFORMATION:
/ APPLICANT: Peterson, David P.
/ APPLICANT: Peterson, Cecelia I.
/ APPLICANT: Cooke, Benjamin G.
/ TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
/ FILE REFERENCE: PA-0029 US
/ CURRENT APPLICATION NUMBER: US/09/971.392
/ CURRENT FILING DATE: 2001-10-03

```

```

? PRIOR APPLICATION NUMBER: 60/737,652
?
? PRIOR FILING DATE: 2000-10-03
?
? NUMBER OF SEQ ID NOS: 260
? SOFTWARE: PERL Program
? SEQ ID NO 65
? LENGTH: 6046
? TYPE: DNA
? ORGANISM: Homo sapiens
?
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: Template ID: 900031.4
?
? US-03-971-392-65

```

Pred. No.:	4,04e-32	Length:	604e6
Score:	432.50	Matches:	106
Percent Similarity:	60.65%	Conservative:	25
Best Local Similarity:	48.07%	Mismatches:	54
Query Match:	23.69%	Indels:	33
GB:	10	Gaps:	2

FIGURE9 (1-361) x US-09-971-392-65 (1-6046)

QY 99 PheAlaIleValAlaLysCys-----LysGLYThrClwMetGluYrLlaAlaLysPhe 116
 Db 551 TTTCCGGTGTGTAAAGAAATGCCCGAGAAAGACACGGCCCTCCAGATACCGGCAATTC 610
 QY 117 IleLysLysArgArgLysProSerSerArgArgLysValSerArgLysArgLysLysArg 136
 Db 611 ATCAAGAAAGAGAGACTAATGCCAGCCGGGGGTGTGAGCCGCGAGCATCGAGCGG 670
 QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
 Db 671 GAGGTGAGATCTCTGAAGAGATCCACACCCCAATGTCTATCACTCGTGAACGAGCTAT 730
 QY 157 GluAsnLysThrAspValLeuIleLeuGluIleValSerGlyGlyGluLeuPheAsp 176
 Db 731 GAGAACAAAGACGAGCTATCTCTATCTTGGAAGCTGTTGACAGTGGCGAGCTTTGAC 790
 QY 177 PheLeuAlaGluLysAspHis**Gln-ArgMetArgProArgSerSerSerArgSse 196
 Db 791 TTTCTAGCTGAAGAGAACTTTTAACGAAAGAGAGCAACTAATTTCTCAACAATTT 850
 QY 196 TrpThrValSerThrTrpCysThrProSerAlaSerArgThrLeuThr***SerArgTr 216
 Db 851 CTTAATGGGTCTTACTCTGCACTGCTCTTCAATGGCCCACTTATGTTAAAGCTGAG 910
 QY 216 GThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
 Db 911 AACATATAGCTTTGGATAGAAATGTGCCCAACCTCGATCAAGTACTATGACTTTGGG 970
 QY 236 AsenArgThrGlySerArgProValAlaSerSerArgThrSerThrLeuAlaArgLysSer 256
 Db 971 T----- 971
 QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyLe 276
 Db 972 -----TGGCCCAATAAATTGACTTTGGAATGATTTTAA 1006
 QY 276 ThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrTr 296
 Db 1007 AACATATATTGGG-ACTCAGAGATTGTGCGTCTCGA-GATAGCAACTATGAACCTTTG 1066
 QY 296 PheGluLyser**TyrValGluHisAlaGlyArgHisHisLeuHisAspPro 311
 Db 1065 GTCCTTAGGACAGATATGTGAGATTCGGGGTAAATAAATCAATATACCT 1110

RESULT 14
 US-10-252-157-452
 ; Sequence 452, Application US/10252157
 ; Publication No. US20030190640A1
 ; GENERAL INFORMATION:
 APPLICANT: Paris, Mary

```
APPLICANT: Pearson, Cecelia I
TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
FILE REFERENCE: PA-0027-1 US
CURRENT APPLICATION NUMBER: US/10/252,157
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: 60/235,048
PRIOR FILING DATE: 2001-05-31
NUMBER OF SEQ ID NOS: 501
SOFTWARE: PERL Program
SEQ ID NO 452
LENGTH: 6046
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030190640A1 900031.4
US-10-252-157-452
```

```
Alignment Scores:
Pred. No.: 4,046-32 Length: 6046
Score: 432.50 Matches: 106
Percent Similarity: 60.65% Conservative: 25
Best Local Similarity: 49.07% Mismatches: 54
Query Match: 23.69% Indels: 33
DB: Gaps: 2
```

FIGURE9 (1-361) x US-10-252-157-452 (1-6046)

```
Qy 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 551 TTTCGGGTTGGTGAAGAATCCCTGAGAGAAAGCCCGCTCCAGTGGTGGCCCAATTC 610
Qy 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
Db 611 ATCAAGAAAGAGAGAGACTAGTCCAGCCGGGGGTGTGAGCCGCGAGACATCGAGCG 670
Qy 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 671 GAGGTGACATCCCTGGAAGAGATCCAGCACCCCAATTCATCACTCCGACGAGGTCTAT 730
Qy 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGluLeuPheAsp 176
Db 731 GAGAACAGAGAGCGATCATCTGATCTTGGAACTCTTGCAGGTGCGAGCTTTGAC 790
Qy 177 PheLeuAlaGluLysAspHis**Gln-ArgMetArgProArgSerSerSerArgSse 196
Db 791 TTCTTACCTGAAAGGAATCTTTAAGTGAAGAGCAACTGAAATTTCTCAACAAT 850
Qy 196 rTyrThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
Db 851 CTATATGTTGTTTCTACCTGACCTCTCTTCAATTCCTCCCACTTTGATCTTAAGCTGAG 910
Qy 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 911 AACATAATGCTTTGATAGATAATGTCCTCCCAACCTCGATCAAGATCATTTGACTTTGGG 970
Qy 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgIleSse 256
Db 971 T----- 971
Qy 256 rSerValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyLe 276
Db 972 -----TGGCCCATTAATAATTGACTTTGGAATAATGATTTAA 1006
Qy 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLeuLeu**ThrThrTr 296
Db 1007 AACATATTGCG-ACCTCGAGAGTTGTGCGCTCGA-GATAGTCAACTATGAACCTTGG 1064
Qy 296 pleuGlySer**TyrValGluHisArgArgHisHisLeuHisPro 311
Db 1065 GTCTTGAGGAGATATGTGAGTATCGGGTAAATACCTATATCT 1110
```

RESULT 15

```
US-10-087-192-1832
Sequence 1832, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1832
LENGTH: 1592
TYPE: DNA
ORGANISM: Mus musculus
US-10-087-192-1832
```

```
Alignment Scores:
Pred. No.: 2,046-31 Length: 1592
Score: 418.00 Matches: 91
Percent Similarity: 75.18% Conservative: 15
Best Local Similarity: 64.54% Mismatches: 32
Query Match: 22.89% Indels: 3
DB: Gaps: 1
```

FIGURE9 (1-361) x US-10-087-192-1832 (1-1592)

```
Qy 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 69 TTTCGCTATGTAAGAATCCCTGAGAGAGAGAGAGAGAGAGAGTATGACGCCAATTC 128
Qy 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
Db 129 ATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 188
Qy 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 189 GAGGTGACATCCCTGGAAGAGATCCAGTGTGTCACCCCAATCATCAGCTGACAGAGCTAT 248
Qy 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGluLeuPheAsp 176
Db 249 GAGAACCCGACAGAGTGTGCTCATCTTGAAGTATGTCGAGAGAGAACTGTTGAT 308
Qy 177 PheLeuAlaGluLysAspHis**Gln-ArgMetArgProArgSerSerSerArgSse 196
Db 309 TTCTTACCTGAAAGGAATCTTTAAGTGAAGAGCAACTGAAATTTCTCAACAAT 368
Qy 196 rTyrThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
Db 369 CTGATGAGGAGTGAATTAATCTTCAACAAAGAAATGCTCTTATGATCTCAAGCCGAA 428
Qy 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 429 AACATCAATGTTTGAACAAGATATCCCATTCACACATCAAGCTGATGACTTTGGC 488
Qy 236 a 236
Db 489 C 489
```

Search completed: July 1, 2004, 01:53:07
Job time : 535 secs

Thu Jul 1 09:50:13 2004

figures.png

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd

OM protein - nucleic search, using frame_plus_p2n model

```
Run on:      July 1, 2004, 01:24:35 ; Search time 483 Seconds
              (without alignments)
              3175.154 Million cell updates/sec
```

Title:	FIGURE9
Perfect score:	1826
Sequence:	1 XMXIGIGRSSPXQLDLS.....GXEMSLGTLMPGPHQXRTX 361

Scoring table:		
	BLOSUM62	
Xgapop	10.0	Xgapext 0.5
Ygapop	10.0	Ygapext 0.5
Zgapop	6.0	Zgapext 7.0
Delcp	6.0	Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

```

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=cg22.1/USTO.spool/GBEBSYBUS41/rnmt.29062004.14350.24548/app.query.fasta_1.5139
-D=N.geneSeq.294904 -OFMT=fastap -SUFFIX=.trng -MINMATH=0.1 -IOGCST=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=-1 -MATRIX=blotsum62 -PARAM=human40.cdi
-LIST=45 -DOCLIN=ln-200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 ALG=0.15
-MODE=IOCL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=GBEBSYBUS41.CSGN.1.1.470 @rnmt.29062004.14350.24548 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DBP=block=100 -LONLOG
-DEV=TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOB=10 -XGAPEXT=0.5 -FGAPOB=6
-FGAPEXT=7 -YGAPOB=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database :
1: geneseqn129Jan04:*
2: geneseqn12906:*
3: geneseqn20006:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1746.5	95.6	1093	6	ABSS4622	Abss4622 Rat myosi
2	865.5	46.9	1514	9	ADBS8445	ADBS8445 Toxicity
3	865.5	46.9	1514	9	ADBS3052	ADBS3052 Primary r
4	732	39.3	1428	2	AAK34657	AAK34657 Korrine Z1
5	632	34.6	2079	4	AAK34281	AAK34281 Human tcl
6	632	34.6	2105	6	ABV72291	ABV72291 Nucleotid
7	632	34.6	2132	2	AAK34656	AAK34656 Human ZIP
8	632	34.6	2132	7	ACA90235	ACA90235 Deatch as

9	632	34.6	2226	9	ADD92770	Adh9p	Human	tum
10	631.5	34.6	2224	5	AAH61518	Aah61518	Human	CDN
11	631.5	34.6	2224	5	AAH78058	Aah78058	Nucleotide	
12	581.5	31.8	727	4	AAK93282	Aak93282	Human	CDN
13	581.5	31.8	757	4	AAK91856	Aak91856	Human	CDN
14	548	30.0	1268	4	AAK70641	Aak70641	Human	Imm
15	469.5	25.7	1253	5	AAK73461	Aak73461	DNA	encod
16	462.5	25.3	1742	3	AAZ49755	Aaz49755	Human	DAR
17	462.5	25.3	1742	3	AAZ49755	Aaz49755	Human	DAR
18	432.5	23.7	992	9	ADD98021	Add98021	Human	deac
19	432.5	23.7	1010	9	ADD98022	Add98022	Human	deac
20	432.5	23.7	5910	6	ABK86815	Abk86815	CDNA	enric
21	432.5	23.7	5910	6	ADD14655	Add14655	Human	src
22	432.5	23.7	6046	9	ADBD47365	Adbd47365	Human	CDN
23	432.5	23.7	6046	9	ADBD4105	Adbd4105	Human	pro
24	428.5	23.5	4272	2	AAQ69839	Aaq69839	Human	deac
25	428.5	23.5	5886	2	AAQ69839	Aaq69839	Human	deac
26	428.5	23.5	5886	2	AAQ69839	Aaq69839	DNA	sequ
27	335.5	18.4	493	8	ACH05087	Ach05087	Human	mem
28	330	18.1	539	8	ACH06178	Ach06178	Human	mem
29	295	16.2	438	8	ACH14044	Ach14044	Human	CDN
30	278.5	15.3	491	4	AA110656	Aa110656	Probe	#57
31	278.5	15.3	491	4	ABAA52263	Abaa52263	Human	Foe
32	278.5	15.3	491	4	AAI11900	Aai11900	Probe	#58
33	278.5	15.3	491	4	ABAA22094	Abaa22094	Probe	#56
34	278.5	15.3	491	4	AAK62017	Aak62017	Human	bon
35	278.5	15.3	491	4	AAK00568	Aak00568	Human	bra
36	278.5	15.3	491	4	ABE32564	Abes2564	Human	lrv
37	278.5	15.3	491	5	AA100576	Aa100576	Probe	#56
38	278.5	15.3	491	6	ABSO00558	Abso00558	Human	gen
39	264	14.5	259	4	AA119895	Aa119895	Probe	#98
40	264	14.5	259	4	ABAA49922	Abaa49922	Human	foe
41	264	14.5	259	4	AA145091	Aa145091	Probe	#13
42	264	14.5	259	4	ABAA47040	Abaa47040	Human	brea
43	264	14.5	259	4	ABAA22024	Abaa22024	Probe	#10
44	264	14.5	259	4	AAK39086	Aak39086	Human	don
45	264	14.5	259	4	AAK13350	Aak13350	Human	bira

ALIGNMENTS

RESULT 1
ABS54622
ID ABS54622 standard; cDNA; 1093 BP.

AC	ABSS4622;
XX	
DT	28-NOV-2002 (first entry)

DE Rat myosin phosphatase targeting subunit 1-kinase, MYPT1 kinase, CDNA.

KM Rat; ss; gene; MYPT1 kinase, myosin phosphatase targeting subunit
KM SMP-1M; myosin light chain phosphatase; smooth muscle disease;
KM inotropic; hypotensive; muscular; hypertension.

OS Rattus sp.

FN	Key	Location/Qualifiers
FT	CDS	1..1092
FT		/*tag= a
FT		/product= "MPT1-kinase"
FT		/note= "No stop codon shown"
FT		/transl_except= (pos:1..3,aa:Xaa)
FT		/transl_except= (pos:7..9,aa:Xaa)
FT		/transl_except= (pos:37..39,aa:Xaa)
FT		/transl_except= (pos:103..105,aa:Xaa)
FT		/transl_except= (pos:172..174,aa:Xaa)
FT		/transl_except= (pos:559..561,aa:Xaa)
FT		/transl_except= (pos:646..648,aa:Xaa)
FT		/transl_except= (pos:814..816,aa:Xaa)
FT		/transl_except= (pos:871..873,aa:Xaa)
FT		/transl_except= (pos:886..888,aa:Xaa)
FT		/transl_except= (pos:907..909,aa:Xaa)

PT /transl_except= (pos:952..954,aa:Xaa)
 PT /transl_except= (pos:1036..1038,aa:Xaa)
 PT /transl_except= (pos:1054..1056,aa:Thr)
 PT /transl_except= (pos:1081..1083,aa:Xaa)
 PT /transl_except= (pos:1090..1092,aa:Xaa)
 PT /note= "Xaa is unknown"

W0200267765-A2.

06-SEP-2002.

27-FEB-2002; 2002W0-US005736.

27-FEB-2001; 2001US-0271436P.

(UYDU-) UNIV DUKE.

Haystead TA;

WPI; 2002-698619/75.

P-PSDB; ABG70856.

New myosin phosphatase targeting kinase, useful for identifying compounds that can treat hypertension or smooth muscle disease.

Claim 6; Fig 8; 33pp; English.

The invention relates to myosin phosphatase targeting subunit 1-kinase, MYPT1 kinase (a subunit of the myosin light chain phosphatase SMP-1M) or a portion of at least 5 consecutive amino acids. Also included are the nucleic acid encoding MYPT1 kinase, a MYPT1 kinase expression vector, a transformed host cell, an anti-MYPT1 kinase antibody, screening a test compound for anti-hypertensive activity and a kit for detecting the MYPT1 kinase comprising a compound that specifically binds to MYPT1 kinase disposed within a container means. MYPT1 kinase is useful for identifying compounds that inhibit hypertensive activity or which treat smooth muscle disease. The present sequence encodes a MYPT1-kinase from rat aorta smooth muscle

Sequence 1093 BP; 223 A; 318 C; 318 G; 227 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.: 9,88e-96 Length: 1093
 Score: 1746.50 Matches: 349
 Percent Similarity: 96.41% Conservative: 0
 Best Local Similarity: 96.41% Mismatches: 10
 Query Match: 95.65% Indels: 3
 Gaps: 2

FIGURE9 (1-361) X ABS54622 (1-1093)

QY 2 Met***IleGIYLeuIleGIYArgSerPro***GIYGILeuAspSerLeuSerAsp 21
 Db 4 ATGATATCGGTTTAATCGCGGAGCTCGCCNCGGAGCTGACCTCCCTCTCAAC 63
 QY 22 LeuLeuLeuSerArgProGlnHisGIYIleAsnLeuThr**LeuPheLeuGIYProArg 41
 Db 64 CTCCTCTCTTCTGCGCCCTCAAGCAGGATTAACCTCACTGACCTGTTGGGCCCCG 123
 QY 42 CysArgIleSerValLeuSerLeuIleValIleProIleCysLeuSer***GlySerLeu 61
 Db 124 TGCAGGCGCAGCGCTCTCTCCCTCAAGCAATCCCAAGGTCTGTATAGAGGCTCTTG 183
 QY 62 GlySerSerValIleValIleAsnLeuGIYThrAspAlaGlnArgLeuGIYTYArgVal 81
 Db 184 GGCAGCTCTGTTGTTGCGGAACCTGGGAACATGACAGAGGCTGGGGTACAGAGTC 243
 QY 82 LeuProSerSerGIYSerAlaIleLeuSerCysSerPheProHisSerGIY---PheAla 100
 Db 244 CTGCTCTCTCTGCGGCTGACAGCGCTTACGTCTCTCCCAAGCGCGCAGTTCTGCC 303
 QY 101 IleValArgIleCys-----LysGIYThrGIYMetGlnTYArgAlaIleValPheIleLys 118

Db 304 ATCGTGGCAAGTGCACAGAGGGAACCGGCGATGAGTACCGGCGCAAGTTCAATAAG 363
 QY 119 LysArgArgLeuProSerSerArgArgGIYValSerArgGlnGlnIleGIYArgVal 138
 Db 364 AAGCGCGCCCTGCGCTCCAGCGCGCGGCTGTAGCCGTGAGAGATCGAGCGGAGAGTG 423
 QY 139 SerIleLeuArgGlnIleArgHisProAsnIleIleThrLeuHisAspValPheGlnAsn 158
 Db 424 AGCATCTCGCGCGAGATCCCGCACCCCAACATCATCTACCGCTGACAGATGTGTGAGAAC 483
 QY 159 LysThrAspValIleValIleLeuGIYLeuValSerGIYGIYLeuPheAspPheLeu 178
 Db 484 AAGACAGATGTGGTCTGATCTTGGAGCTGTGTCGGCGCGAATCTTTCGATTTCTG 543
 QY 179 AlaGlnIysAspHis**GlnArgMetArgProArgSerSerSerSerArgSerTyrThr 198
 Db 544 GCTGAGAGATCACTGACAGAGATGAGCGCAGCGCATCTCTCAAGCATCTCTGACG 603
 QY 199 ValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgArgThrSer 218
 Db 604 GTGTCACTACCTGACTCCAAAGCGCATCGGCACTTTGACCTGAAAGCCGAGAACATCA 663
 QY 219 CysCysTyrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAlaSerArg 238
 Db 664 TCTTCTGAGACAGATGATCAGCAGCCAGCCATTAAGCTCATGACTTTGGCATCGCGC 723
 QY 239 ThrGIYSerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSerVal 258
 Db 724 ACGAGATGAGCGCGGTGACAGATTCAGAAATCTTTGGACGCCAGAGTTGTCGTG 783
 QY 259 ArgGIYArgCysGIYHisHisProIleGIY**IleLeuHisGIYLeuGIYLeuThrCys 278
 Db 784 AGGCGCAGATGTGGGCAACACCCGATAGGAGTATTTGGACGCGCTTGCGCTGACCTGC 843
 QY 279 LeuAsnAspProValPheHisSerPro**AspCysIleLeu**ThrThrTyrLeuGIY 298
 Db 844 CTCACAAATCTCTCTTCCACAGCGCCCGAATTTGTAACATGACCACTTGCTTGGA 903
 QY 299 Ser***TyrValGlnHisArgArgHisIleLeuHisProValSerAla**AspGIYGln 318
 Db 904 AGCTGATGTGGAGATCGCGCGTCACTATCACTCTGAGATGCTTGAGATGGGAG 963
 QY 319 GIYProGlnThrValProAlaArgGIYProGIYIleArgAlaGIYThrSerAlaAsnCys 338
 Db 964 GGGCTTCAGACTGTACTGTCTAGAGGCCAGAGATCAGGGCTGACCTCTGCAAAACTGC 1023
 QY 339 LysHisTyrGIY**GlnMetSerLeuGIYThrLeuAspMetProGIYProHisGln** 358
 Db 1024 AAACACTGCGGCTGAGAGATGCTCTGGGAACNCTGATATGCTTGAGCCCAACANGT 1083
 QY 359 ArgThr 360
 Db 1084 AGGACC 1089

RESULT 2
 ADBS8463
 ID ADBS8463 standard; DNA; 1514 BP.
 XX
 AC ADBS8463;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3489.
 XX
 KM Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 XX drug screening; toxicity assay; ds.
 OS Unidentified.
 XX
 PN W02003064624-A2.
 XX
 PD 07-AUG-2003.

```

PF 31-JAN-2003; 2003WO-US003194.
XX
XX 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX
PA (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX WPI; 2003-689530/65.
XX
PT Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
PT the compound.
XX
XX Claim 1; SEQ ID NO 3489; 1156bp; English.
XX
CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 1514 BP; 303 A; 449 C; 514 G; 248 T; 0 U; 0 Other:
SQ
Alignment Scores:
Pred. No.: 2.11e-42 Length: 1514
Score: 856.50 Matches: 188
Percent Similarity: 75.10% Conservative: 2
Best Local Similarity: 74.31% Mismatches: 16
Query Match: 46.91% Indels: 47
DB: 9 Gaps: 4
FIGURE9 (1-361) x ADB58463 (1-1514)
QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 183 TTGCGCATCTGTCGCAAGTCCAGAGAGGCAAGCGGAGTGAAGTACGGCGCAAGTTC 242
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 243 ATAAAGAGAGGGGCGCTGCGTCCAGCCGGCGGGGTGTGAGCGGTAGAGAGATCGAGCC 302
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 303 GAGGTAGGATCTCTGGCGGAGATCCGCCACCCCAACATATCATCGTGCACATGTGTC 362
QY 157 GluAsnLysThrAspValValIleuIleLeuGluLeuValSerGlyGluLeuPheAsp 176
DB 363 GAGAAACAAGACATGTGTGCTGATCTTGAGCTGTGTGCGCGCGCAACTTTTCGAC 422
QY 177 PheLeuAlaGluLysAsp-His**GlnArgMetArgProArgSerSerSerSerArgSe 196
DB 423 TTCTGCGCTAGAGAGAGTACTGACAGAGAGATGAGGCCACGAGTTCTCTCAAGCAGATC 482
QY 196 rTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
DB 483 CTGAGAGGTGTCCACTACCTGACCTCAAGGCGATCGCGCATTTGACCTTGAAGCGGAG 542

```

```

QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 543 AACATCATCTTCTCTGACAGACATGACAGCCAGCCACCATTAACCTCATGCACCTTGGC 602
QY 236 AserArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 603 ATCGGCGACAGATGAGAGCGCGGTAGCGAGTTCAAGAACATCTTTGGCACCGCAGAGTTC 662
QY 256 rSerValArgGlyArgCysGlyHisIleProIleGly**IleLeuHisGlyLeuGlyLe 276
DB 663 G----- 663
QY 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
DB 664 -----TCGCCCTGAGATGTGAACCTTGAACATGAAACCACTT 695
QY 296 rPLeuGlySer**TyrValGluHisArgArgHisIleLeuHisProValSerAla**A 316
DB 696 GCGTTGAGAGCTGATATGTGAGCATCGCGCTCATCACTCATCT----- 742
QY 316 SPGLyGlnGlyProGlnThrValProAlaArgGlyProGlyIleArgAlaGlyThrSerA 336
DB 743 -----CCTGAGCGGCGCTTCCCATCTCTGGGAGACCAAG 779
QY 336 lAsn-----CysLysHisTrpGly**GluMet 345
DB 780 CAGGAACGCTGACAAACATCTCGGCTGTGAACTA 814
RESULT 3
ADBS3052
ID ADB53052 standard; DNA; 1514 BP.
XX
XX ADB53052;
AC
XX 04-DEC-2003 (first entry)
DT
XX
XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3594.
DE
XX
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
XX Rattus norvegicus.
OS
XX
XX WO2003065993-A2.
PN
XX
XX 14-AUG-2003.
PD
XX
XX 04-FEB-2003; 2003WO-US003482.
PF
XX
XX 04-FEB-2002; 2002US-0353171P.
XX
XX 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370246P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 11-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374135P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;

```

PI Elashoff M;
 XX
 DR WPI; 2003-731472/69.
 XX
 PT Determining if a compound induces a toxic effect on a tissue or cell, for
 PT identifying hepatotoxic compounds, comprises comparing a gene expression
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
 PT mean values.
 XX
 PS Claim 44; SEQ ID NO 3594; 874bp; English.
 XX
 CC The present invention describes a method for determining whether a
 CC compound induces a toxic effect on a tissue or cell. The method comprises
 CC preparing a gene expression profile of a tissue or cell sample exposed to
 CC the compound, and comparing the gene expression profile to a database
 CC comprising data or information on the Tox mean and non-Tox mean value.
 CC The method is useful for predicting or identifying at least one toxic
 CC effect, particularly hepatotoxicity, of a test or unknown compound. The
 CC genes listed in the specification are useful as diagnostic or toxicity
 CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. These
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.
 XX
 SO Sequence 1514 BP; 303 A; 449 C; 514 G; 248 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,116-42 Length: 1514
 Score: 856.50 Matches: 188
 Percent Similarity: 75.10% Conservative: 2
 Best Local Similarity: 74.31% Mismatches: 16
 Query Match: 46.91% Indels: 47
 DB: 9 Gaps: 4

FIGURE9 (1-361) x ADBS3052 (1-1514)

```

QY 99 PhealalalevalAArglysCys-----LysGlyThirGlymetGlyrYAlaAlaLysPhe 116
DB 183 TTCCGCATCGGCGCAAGTGCACAGAGAGGACCGGATGAGTACGGCGCCAAAGTTTC 242
QY 117 IlelyslsArGArGleuProSerSerArGlyValSerArGlyGluIleGluArg 136
DB 243 ATTAAGAGCGCGCCCGCCCTCCAGCGCGCGGTGTGAGCGGTGAGAGATCGAGCGC 302
QY 137 GluValSerIleLeuArGlyLueArgHisProAsnIleIleThirleuHisAspValPhe 156
DB 303 GAGGTGACATCTCGCGGAGATCCGCCACCCCAACATCATACCGTGCACGATGTGTTC 362
QY 157 GluAsnlyStrAspValValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 363 GAGAAACAAGACAGATGTGTGCTGATCTTGAGCTGTGTGTCGGCGGCAACTTTTCGAC 422
QY 177 PheleuAlaGlyLysAsp-His***GlnArgMetArProArgSerSerSerSerArgSe 196
DB 423 TTTCGTGCTGAGAGAGATCTCTACACAGAGATAGAGCCACGAGTTCTCTACAGCATC 482
QY 196 TrpThrValSerThrThrCysThrProSerAlaSerArGlyThrIleuThr***SerArG 216
DB 483 CTGACGGGTGTCCACTACCTGACCTCCACAGCGCATCCGCGCACTTGAACCTGAAGCCCGAG 542
QY 216 gThnSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrIleuAl 236
DB 543 AACATCATGTTGCTGAGACAGATGACAGCCAGCCCACTTAAGCTCATGCACTTGGC 602
QY 236 aserArGThrIleSerArGProValAlaSerSerArGThrSerleuAlaArgIleSerSe 256
DB 603 ATCCGCGACAGATGAGAGCGCGGTAGCGAGTTCAAGAAACATCTTGCACGCCAGATTC 662
QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleleuHisGlyLeuGlyLe 276
DB 663 G----- 663
  
```

```

QY 276 UThrCysleuAsnAsnProValPheHisSerPro***AspCysLySleu***ThrThrT 296
DB 664 -----TCGCCCCCTGAGATTTGTAACCTATGAAACCACTT 695
QY 296 rPleuGlySer***TyrValGluHisArGArgHisHisSleuHisProValSerAla***A 316
DB 696 GCGTTGAGAGCTGATATGTGAGCATCGCGCTCATCATCACTACATCCCT----- 742
QY 316 spGlyGlnGlyProGlnThrValProAlaArgGlyProGlyIleArgAlaGlyThrSerA 336
DB 743 -----CCTGACGGGGGCTTCCCATTCCTCGGGGAGACCAAG 779
QY 336 laAsn-----CysLyHisTrpGly***GluMet 345
DB 780 CAGGAAAGCTGACAAACATCTCGCTGTGACTA 814
RESULT 4
AAK34657
ID AAK34657 standard; DNA; 1429 BP.
AC AAK34657;
DT 01-JUL-1999 (first entry)
DE Murine ZIP-kinase (serine/threonine kinase) encoding DNA.
KW Zipper Interacting Protein Kinase; ZIP-kinase; serine/threonine kinase;
KW leucine zipper domain; transcription factor ATF4; gene therapy; cancer;
KW Human; murine; ss.
XX Mus musculus.
XX EP911408-A2.
XX 28-APR-1999.
XX 24-SEP-1998; 98EP-00307747.
XX 26-SEP-1997; 97JP-00261589.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI Akira S. Kawai T;
XX WPI; 1999-246420/21.
DR P-PSDB; AAY06922.
PT New Recombinant Zipper Interacting Protein Kinase (ZIP-kinase) protein
PT and DNA, useful as anticancer agents.
XX
PS Claim 6; Page 19-22; 33pp; English.
XX
CC The invention provides human and murine recombinant Zipper Interacting
CC Protein Kinase (ZIP-kinase) proteins. These proteins are serine/threonine
CC kinases which bind the leucine zipper domain of transcription factor
CC ATF4. Host cells containing vectors comprising the ZIP-kinase nucleic
CC acids are used for the recombinant expression of the proteins. ZIP-kinase
CC protein and DNA are useful as gene therapeutic agents against cancer, and
CC as anti-cancer agents. The present sequence represents a DNA encoding a
CC murine ZIP kinase protein
XX
SQ Sequence 1429 BP; 309 A; 425 C; 475 G; 220 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,156-34 Length: 1429
Score: 722.00 Matches: 163
Percent Similarity: 73.82% Conservative: 9
Best Local Similarity: 69.96% Mismatches: 21
Query Match: 39.54% Indels: 40
DB: 2 Gaps: 5
  
```

FIGURE9 (1-361) x AAK34657 (1-1429)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 79 TTGGCATGCGCGAAGTGGACAGAGGCGCGGACATGAGTATGACGCCAAGTTCC 138
QY 117 IleuYslyArgArgLeuProSerSerArgArglyValSerArgGlnGluIleGluArg 136
Db 139 ATCAAGAAAGCGCGCGCTGCCATCCAGCGCGCGGTGTGAGCCGGAGAGATGAAAGC 198
QY 137 GluValSerIleLeuAlaGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 199 GAGGTGAGCATCCCGCGAGATCCGCCACCCCAACATCATTAACATCGATGACGTGTTC 258
QY 157 GluAsnLysThrAspValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 259 GAGAACAAAGACAGATGCTGCTCATCTCTGAGAGTGTGTCGGTGGCGAGCTTTTGCAC 318
QY 177 PheLeuAlaGluLysAsp-His***GlnArgMetArgProArgSerSerSerArgSe 196
Db 319 TTCTGGCGCGAGAGAGTCAATGACGAGAGATGAGCGACAGCTTCTCAAAACAATC 378
QY 196 rTrrPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 379 CTAAACCGCTGCTCTACCTGCTGCACTCCAAAGCGATCCCAACTTTCGTAAGCCGAG 438
QY 216 gThrSerCysSerTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 439 AACATCATGTTGTTGGACAAAGACGACGCCGCCCGCATTAAGCTCATCGACTTTGGC 498
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 499 ATCCGCGACAGGACGAGCTGCGAGCTGCGAGAGTTCAGAAACAATCTTTGCCA----- 547
QY 256 rSerValArgGlyArgCysGlyHisHisPro-IleGly***IleLeuHisGlyLeuGlyL 276
Db 548 -----CACCGAGTT----- 557
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrT 296
Db 558 ----TGT-----CGCCCCGAGATCGTGAACATATGAGCCACTT 591
QY 296 rPLeuGlySer***IyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316
Db 592 GCGCTGAGGCTGACATGTGAGCATTTGGCGTCACTACCTCAATCCTCTCTGAG----- 644
QY 316 spGlyGlnGlyProGlnThrValProAlaArgGly 327
Db 645 -----CGAGCGTCCCGCATCTCTGGG 665

```

RESULT 5

```

AAK94258
ID AAK94258 standard; cDNA; 2079 BP.
XX AAK94258;
AC AAK94258;
XX 06-NOV-2001 (first entry)
DE Human full-length cDNA, SEQ ID NO: 2874.
XX Human full-length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
OS Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-00114089.
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX

```

```

PA (HELI-) HELIX RES INST.
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR P-PSDB; AAK93338.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
PS Claim 8; SEQ ID NO 2874; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO
XX
SQ Sequence 2079 BP; 396 A; 626 C; 733 G; 324 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	Length:	2079
Score: 7.71e-29	Matches: 155	
Percent Similarity: 632.00	Conservative: 13	
Best Local Similarity: 57.69%	Mismatches: 46	
Query Match: 34.61%	Indels: 54	
DB: 4	Gaps: 9	

FIGURE9 (1-361) x AAK94258 (1-2079)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
Db 139 TTGGCATGCTGCGAAGTGGACAGAGGCGCGGACAGAGATGACGCCAAGTTCC 198
QY 117 IleuYslyArgArgLeuProSerSerArgArglyValSerArgGlnGluIleGluArg 136
Db 199 ATCAAGAAAGCGCGCGCTGCTCAATCCAGCGCGGTGAGCGGAGAGATCGAGCGG 258
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 259 GAGGTGAGCATCCCGGAGATCCGGAACCCCAACATCATCCTTCCAGACATCTTC 318
QY 157 GluAsnLysThrAspValIleuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 319 GAGAACAAAGACAGATGCTGCTCATCTCTGAGAGTGTGTCGGTGGCGAGCTTTTGCAC 378
QY 177 PheLeuAlaGluLysAspHis***GlnArgMetArgProArgSerSerSerArgSe 196
Db 379 TTCTGGCGCGAGAGATCCCGAGATCCGCCACCCCAACATCATTAACATCGATGACGTGTTC 438
QY 196 rTrrPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 439 CTGAGCGCGCTTCACTACCTGACCTCAAGCGATCCGACACTTTCAGAGCCGGA 498
QY 216 gThrSerCysSerTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 499 AACATCATGCTGCTGAGAACAGCTGCCAACCCCAACATCAAGCTCATGACTTCGCG 558
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 559 ATCCGCGACAGATGAGAGCGCGGAGACAGATTCAGAAACA-----TC 600
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
Db 601 TTGGCACCCCGGAGTTGTGTC----- 623

```

```

QY 276 eutHrCysLeuAsnProValPheHisSerPro**AspCysLysLeu**ThrThr 296
Db 624 -----CCAGAGATTGTGAACATATGAGCCGCG 651
QY 296 rpleuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
Db 652 GGCCTGAGCGCGACATGTGAGCATGTGTATCATCCTATATCTCT----- 698
QY 316 spgLyGlnGlyProGln-----ThrValProAlaArgGlyProGlyIleArgAlaG 333
Db 699 -----CCTGAGCGGTGCATCCCCGTCTCTGGCCAGAC-----CAAGCAG 738
QY 333 lYThrSerAlaAsnCysLysHisTyrGly**GluMetSerLeu-----GlyThrLeuA 351
Db 739 GAGAGCGCTCAC-----CAACATCTCAGCCGTGAACTAGCACTTCGACGAGAGTACTTC 792
QY 351 spMetProGlyProHisGln 357
Db 793 AGCAA-----CACCGAG 803

RESULT 6
ABV72291
ID ABV72291 standard; DNA; 2105 BP.
AC ABV72291;
XX 16-DEC-2002 (first entry)
XX Nucleotide sequence of human DAPK3.
XX
XX Human; IFIT-2; chronic myelogenous leukemia; LAGE-1; BAGE; DDB1; ETS2;
XX P1ASX; P1ASX-alpha; P1ASX-beta; DAPK3; gene; ss.
XX
XX Homo sapiens.
XX
XX Key location/Qualifiers
XX FH 94..1458
XX FT CDS /*tag=a
XX FT /product= "DAPK3"
XX
XX WO200270747-A1.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002MO-JP001901.
XX
XX 01-MAR-2001; 2001JP-00056438.
XX
XX (FUJI ) FUJISAWA PHARM CO LTD.
XX
XX Mano H;
XX
XX WPI; 2002-682911/73.
XX
XX P-PSDB; ABB78353.
XX
XX Measuring the expression profile of genes in a cell or tissue sample for
XX diagnosis of chronic myelogenous leukemia and identification of agents
XX for its treatment.
XX
XX Disclosure; Page 81-85; 97pp; Japanese.
XX
XX The present sequence encodes human DAPK3. The expression level of the
XX gene is used in the method of the invention. The specification describes
XX a method of examining chronic myelogenous leukemia. The method comprises
XX measuring the expression level of a gene selected from IFIT-2, LAGE-1,
XX BAGE, DDB1, ETS2, P1ASX, P1ASX-alpha, P1ASX-beta and DAPK3, or
XX determining the expression profile of a group of genes including one or
XX more of these genes, in a cell or tissue sample from a chronic
XX myelogenous leukemia patient. The method is used for the diagnosis,
XX treatment and prevention of chronic myelogenous leukemia
XX
XX Sequence 2105 BP; 402 A; 630 C; 741 G; 332 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 7,81e-29 Length: 2105
Score: 632.00 Matches: 155
Percent Similarity: 62.69% Conservative: 13
Best Local Similarity: 57.84% Mismatches: 46
Query Match: 34.61% Indels: 54
Db: 6 Gaps: 9

FIGURE9 (1-361) x ABV72291 (1-2105)
QY 99 PheAlaIleValIAsCysCys-----LysGlyThrGlyMetGlyTyrAlaAlaLysPhe 116
Db 163 TTGGGATCTGTCGGAAATGCGCGAGAGGCGACAGGCAAGAGTACGCGCCAGTTC 222
QY 117 lIeLysLysArgArgLeuProSerSerArgArgGlyValSerArgGlnGlnIleGlyArg 136
Db 223 ATCAAGAAAGCCCGCTGTCTATCCAGCCGCGGTGGGGTGAGACCGGAGAGATGACCGG 282
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 283 GAGGTGAACATCTCTCGGGAATCCGGAATCCGCAACCAATCATCATCCCTGACATCATCTC 342
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
Db 343 GAGAACAAAGACGAGCATGTGCTCTCATCTCGAGAGCTGTCTCTGCGGGAGACTCTTTGAC 402
QY 177 PheLeuAlaGluLysAspHis***GlnArgMetArgProArgSerSerSerArgSer 196
Db 403 TTCTCGGCGAGAAAGATCGCTGACGAGAGACGAGCCACCCAGTTCCTCAACGACATC 462
QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 463 CTGAGCGGGTTCATCTACTGACTGACTCTTAAGCGCATCGCACCTTGACCTGAACCGGAA 522
QY 216 rHisSerCysCysTyrThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 523 AACATCATGCTGCTGTGACCAAGACCTGCCAACCACCAATCATGACTCATCTTGGC 582
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 583 ATCGCGCACAAATGAGACGGCGGGAAGATTCAAGAAC-----TC 624
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
Db 625 TTGCGCACCCCGAGCTTGTGTC----- 647
QY 276 eutHrCysLeuAsnProValPheHisSerPro**AspCysLysLeu**ThrThr 296
Db 648 -----CCAGAGATTGTGAACATATGAGCCGCGT 675
QY 296 rpleuGlySer**TyrValGluHisArgArgHisHisLeuHisProValSerAla**A 316
Db 676 GGCCTGAGCGCGACATGTGAGCATGTGTATCATCCTATATCTCT----- 722
QY 316 spGlyGlnGlyProGln-----ThrValProAlaArgGlyProGlyIleArgAlaG 333
Db 723 -----CCTGAGCGGTGCATCCCCGTCTCTGGCCAGAC-----CAAGCAG 762
QY 333 lYThrSerAlaAsnCysLysHisTyrGly**GluMetSerLeu-----GlyThrLeuA 351
Db 763 GAGAGCGCTCAC-----CAACATCTCAGCCGTGAACTAGCACTTCGACGAGAGTACTTC 816
QY 351 spMetProGlyProHisGln 357
Db 817 AGCAA-----CACCGAG 827

RESULT 7
AAK34656
ID AAK34656 standard; DNA; 2132 BP.
XX
XX AAK34656;

```



```

DT 01-JUL-1999 (first entry)
XX Human ZIP-kinase (serine/threonine kinase) encoding DNA.
DE
XX Zipper Interacting Protein Kinase; ZIP-kinase; serine/threonine kinase;
XX leucine zipper domain; transcription factor Arf4; gene therapy; cancer;
KW Human; murine; ss.
XX
XX Homo sapiens.
XX
XX EP911408-A2.
XX
XX 28-APR-1999.
XX
XX 24-SEP-1998; 98EP-00307747.
XX
XX 26-SEP-1997; 97JP-00261589.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Akira S, Kawai T;
XX
XX WPI, 1999-246420/21.
XX P-PDB; AAY06921.
XX
XX New Recombinant Zipper Interacting Protein Kinase (ZIP-kinase) protein
XX and DNA, useful as anticancer agents.
XX
XX Claim 5, Page 15-18; 33pp; English.
XX
XX The invention provides human and murine recombinant Zipper Interacting
XX Protein Kinase (ZIP-kinase) proteins. These proteins are serine/threonine
XX kinases which bind the leucine zipper domain of transcription factor
XX ATF4. Host cells containing vectors comprising the ZIP-kinase nucleic
XX acids are used for the recombinant expression of the proteins. ZIP-kinase
XX protein and DNA are useful as gene therapeutic agents against cancer, and
XX as anti-cancer agents. The present sequence represents a DNA encoding a
XX human ZIP kinase protein
XX
XX
SQ Sequence 2132 BP; 429 A; 630 C; 741 G; 332 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,92e-29 Length: 2132
Score: 632.00 Matches: 155
Percent Similarity: 62.69% Conservative: 13
Best Local Similarity: 57.84% Mismatches: 46
Query Match: 34.61% Indels: 54
DB: 2 Gaps: 9

FIGURE 9 (1-361) x AAX34656 (1-2132)
QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluIuYzAlaIaLysPhe 116
DB 163 TTGGCATGCGTGGGAAGTGC CGGAGAGGGGACGGGACAGAGAGTACCAAGCAAGTTC 222
QY 117 IleLysLysArgArgLeuProSerSerArgArgGlyValSerArgGluGluIleGluArg 136
DB 223 ATCAAAAGCGCGCGTGCATCCACCGCGGCGGTAGCGGGAGAGAGATCAAGCGG 282
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 283 GAGGTAAACATCCTCGGGAGATCCGGACCCCAACATATATACCTTGACACATCTTC 342
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyLysLeuPheAsp 176
DB 343 GAGAAACAAGACGACGCTGCTCATCTCGAGCTGCTCTGGCGGGAGCTCTTTGAC 402
QY 177 PheLeuAlaGluLysAspHis***GlnArgMetArgProArgSerSerSerArgSer 196
DB 403 TTCCTGCGGAGAAAGTCCGCTGACGAGAGACAGAGCCACCGATCTCTCAAGCAGATC 462
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216

```

```

DB 463 CTGACGCGGCTCACTACCTGACATCTAAGCGCATCGACACTTGACCTGAAGCCGGA 522
QY 216 gThrSerCysCysTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 236
DB 523 AACATCATCTCTCTGAGCAAGACGTCGCCAACCCACCAATCTCAAGCTCATGACTTCGGC 582
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 583 ATCGGCGCAACAGATGAGCGCGGAGAACGAGTTCAAGACA-----TC 624
QY 256 rSerVal ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
DB 625 TTCGGACCCCGGAGTTGTGGC----- 647
QY 276 euThrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThr 296
DB 648 -----CCAGAGATTGTGAATATGAGCGCGCTG 675
QY 296 rPLeuGlySer***TyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316
DB 676 GGCCTGAGCGCGGACATGTGAGCATCGTGTGCATCACCTATATCT----- 722
QY 316 sPGLyGlnGlyProGln-----ThrValProAlaArgGlyProGlyIleArgAlaG 333
DB 723 -----CTGAGCGGTGCATCCCGCTTCCTGGCGGAGAC-----CAAGCAG 762
QY 333 lYThrSerAlaAsnCysLysHisTyrGly***GluMetSerLeu-----GlyThrLeu 351
DB 763 GAGACGCTCAC-----CAACATCTCAGCGGTGAACCTACGACTTCGACGAGAGTACTTC 816
QY 351 sPMeTProGlyLProHisGln 357
DB 817 AGCAA-----CACCG 827
QY
DB
RESULT 8
ACB90235
ID ACA90235 standard; cDNA; 2132 BP.
XX
XX ACA90235;
XX
XX 10-JUL-2003 (first entry)
XX
XX Deatch associated protein kinase 3 (DAPK3) interacting cDNA #7.
XX
XX Human; cytosolic; DAPK3-Agonist; DAPK3-Antagonist; cancer;
XX death-associated protein kinase 3 interacting protein;
XX DAPK3 interacting protein; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003031571-A2.
XX
XX 17-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031357.
XX
XX 05-OCT-2001; 2001US-0327454P.
XX
XX 09-OCT-2001; 2001US-0327917P.
XX
XX 09-OCT-2001; 2001US-0328029P.
XX
XX 09-OCT-2001; 2001US-0328056P.
XX
XX 12-OCT-2001; 2001US-032849P.
XX
XX 15-OCT-2001; 2001US-0329414P.
XX
XX 17-OCT-2001; 2001US-0330142P.
XX
XX 22-OCT-2001; 2001US-0341058P.
XX
XX 24-OCT-2001; 2001US-0343629P.
XX
XX 29-OCT-2001; 2001US-0349575P.
XX
XX 01-NOV-2001; 2001US-0346357P.
XX
XX 25-JUN-2002; 2002US-0351342P.
XX
XX 01-OCT-2002; 2002US-00262445.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Allobrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;

```

PI Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R,
PI Mezes PS, Millet I, Ooi CE, Patrujan M, Rieger DK, Spytek KA,
PI Taupier RJ, Zernhusen BD, Zhong H, Zhong M,
XX WPI: 2003-381704/36.
DR P-PSDB; ABUS9819.

XX New DAPK3 polypeptide, useful for preparing a composition for treating or
PT preventing e.g., cancer.
XX

XX Example 20D: Page 234-235; 253pp; English.

XX The invention describes an isolated polypeptide comprising any of 33 90-
CC 1273 amino acid sequences (I) given in the specification or its mature
CC form, a sequence that is at least 95 % identical to (I), or a sequence
CC comprising one or more conservative substitutions in the amino acid
CC sequence of (I). The polypeptide is useful for preparing a composition
CC for treating or preventing e.g., cancer. This sequence encodes a death
CC associated protein kinase 3 (DAPK3) interacting protein associated with
CC the identification of novel human proteins and their functions

XX Sequence 2132 BP; 429 A; 630 C; 741 G; 332 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.92e-29	Length:	2132
Score:	632.00	Matches:	155
Percent Similarity:	62.69%	Conservative:	13
Best Local Similarity:	57.84%	Mismatches:	46
Query Match:	34.61%	Indels:	54
DB:	7	Gaps:	9

FIGURE9 (1-361) x ACA90235 (1-2132)

```

QY 99 PheAlaIleValAlaGlyCys-----LyGlyThrGlyMetGluTrpAlaAlaLysPhe 116
Db 163 TTTCGGATCGTCGGAAGTCCGGCAGAGGCGCAGGCGCAAGAGTACGACCAAGTTTC 222
QY 117 IleLysAspArgLeuProSerSerArgGlyValSerArgGluGluLeuArg 136
Db 223 ATCAAGAGCGCCGCTGTATCATCAGCGGGTGTAGCGGCGAGAGATCGACGG 282
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
Db 283 GAGGTGAACATCCCGCGAGATCCGGCACCCCAACATCATCACCCTGCACGACATCTTC 342
QY 157 GluAsnLysThrAspValIleLeuIleLeuGluLeuValSerGlyGluLeuPheAsp 176
Db 343 GAGAACAGACGACGATGCTCTCATCTGAGAGCTGTCTCTGCGGCGGAGCTTTGAC 402
QY 177 PheLeuAlaGluLysAspHis***GlnArgMetArgProArgSerSerSerArgSe 196
Db 403 TTCCGGCGGGAAGAAAGTCGTCGACGAGAGCGAGGCAACCCAGTTCTTCACAGCATC 462
QY 196 rTyrThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
Db 463 CTGACGGCGGTTCACTACTCACTCACTTAAGCGCATTCGACACTTTGACCTTAAGCCGGA 522
QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 523 AACATCATGCTGCTGACACAGACAGTGGCCCAACCCAGCATCACTCATGACTTGGC 582
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 583 ATCGGCAACAAGATCGAGCGGCGGACAGATTCAGAACCA-----TC 624
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGlyLeuGlyL 276
Db 625 TTCGGACACCCCGAATTGTGGC----- 647
QY 276 eutThrCysLeuAsnAsnProValPheHisSerPro***AspCysAlaLysLeu***ThrThr 296
Db 648 -----CCAGAGATTGTGAAGTGAAGTGAAGCCGCTG 675

```

```

QY 296 rpleuGlySer***TyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316
Db 676 GGCCTGAGCGGACACATGTCGAGCATGCGTGTATACCTATACCTATACCTATACCT 722
QY 316 spGlyGlnGlyProGln-----ThrValProAlaArgGlyProGlyIleArgAlaG 333
Db 723 -----CCTGACGGGTGATCCCGCTTCTCTGGCGAGAC-----CAAGCAG 762
QY 333 lYThrSerAlaAsnCysAlaHisTrpGly**GluMetSerLeu-----GlyThrLeuA 351
Db 763 GAGACGCTCAC-----CAACATCTCAGCCGTGAACACTACGACTTCAGAGAGTACTTC 816
QY 351 spMetProGlyProHisGln 357
Db 817 AGCAA-----CACCAG 827
QY
Db
RESULT 9
ADD29770
ID ADD29770 standard; mRNA; 2226 BP.
AC ADD29770;
XX 15-JAN-2004 (first entry)
XX
XX Human tumour suppressor mRNA SEQ ID NO:228.
XX ss; human; tumour suppressor; cancer; cancer; cytostatic; gene therapy.
XX Homo sapiens.
XX MO2003058201-A2.
XX 17-JUL-2003.
XX 31-DEC-2002; 2002MO-US041825.
XX 31-DEC-2001; 2001US-0345317P.
XX (QUAR-) QUARK BIOTECH INC.
XX (CLEV-) CLEVELAND CLINIC FOUND.
XX Feinstein B, Gudkov AV;
XX WPI: 2003-598393/56.
XX
XX Diagnosing cancer comprises determining the polypeptide or polynucleotide
XX levels e.g., hepatic lipase, in a sample from a subject, where a higher
XX level compared to that in a subject free of cancer is indicative of
XX cancer.
XX
XX Disclosure; SEQ ID NO 228; 2722p; English.
XX
XX The invention relates to a novel method for diagnosing a cancer in a
XX subject, the method comprises determining, in a sample from the subject,
XX the level of at least one polypeptide, where a higher level of the
XX polypeptide compared to the level of the polypeptide in a subject free of
XX cancer is indicative of cancer. The polypeptide is selected from any of
XX the polypeptides encoded by the polynucleotides listed in the
XX specification and polypeptides which are at least 70% homologous to the
XX polypeptides. The method of the invention has cytostatic activity, and
XX may have a use in gene therapy. The method is useful in identifying
XX markers specific for one or several types of cancer, depending on the
XX tissue origin, which may be used in numerous diagnostic and prognostic
XX applications as well as cancer type-specific targets for therapeutic
XX intervention. The compounds that modulate the activity of a tumour
XX suppressor gene are useful in the treatment of cancer or as anti-cancer
XX drugs. The present sequence represents a polynucleotide of the invention.
XX
XX Sequence 2226 BP; 421 A; 655 C; 804 G; 346 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 8.31e-29 Length: 2226
XX Score: 632.00 Matches: 155

```

Percent Similarity: 62.69%
 Best Local Similarity: 57.84%
 Query Match: 34.61%
 DB: 9
 Gaps: 9

FIGURE9 (1-361) x ADD29770 (1-2226)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 284 TTTCGATCGTGGCGAAGTCCGGCAGAGGCGACGGGCAAGAGTACGAGCCAACTTC 343
QY 117 IleLysLysArgAGLeuProSerSerArgArgGlyAlaSerArgGluGluIleGluArg 136
DB 344 ATCAAGAGAGCGCCCGCTGATCCAGCGCGGTGGGTGAGCCGGGAGAGATCGAGCGG 403
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 404 GAGGTGAACATCCGCGCGGAGATCCGGCAACCCAACTTCACCTTCGACGACATCTTC 463
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 464 GAGAACAGACGAGCGGTCTCATCTCATCTGAGCTGCTCTGCGGAGAGCTCTTGAC 523
QY 177 PheLeuAlaGluLysAspHis***GlnArgMetArgProArgSerSerSerSerArg 196
DB 524 TTCTGCGCGAGAGAGAGTGGCTGACGAGAGAGAGGCGACCCAGATTCCTCAAGCAGATC 583
QY 196 rTTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 584 CTGAGACGCGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 643
QY 216 GlnThrSerCysCysThrThrThrSerMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236
DB 644 AACATCTGCTGCTGCGACAGAAAGTGGCCCAACCCAGAAATCAAGCTTCAGCTTCGCGG 703
QY 236 aserArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 704 ATGGCGCAACAATCGAGCGCGGAGAGAGTTCAGAGAC-----TC 745
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
DB 746 TTGCGGACCCCGGAGATTGTGGC----- 768
QY 276 euthrCysLeuAsnAsnProValPheHisSerPro***AspCysLysLeu***ThrThrT 296
DB 769 -----CCGAGAGATTGTGAATGTAGCCGCTG 796
QY 296 rPLeuGlySer***TyrValGluHisArgArgHisHisLeuHisProValSerAla***A 316
DB 797 GGCTGTGAGGCGGAGCATGTGAGCATCGGTGCATCACCCTATATCT----- 843
QY 316 spGlyGlnGlyProGln-----ThrValProAlaArgGlyProGlyIleArgAlaG 333
DB 844 -----CTGAGCGGTGCATCCCGCTCTCTGCGAGAC-----CAAGCAG 883
QY 333 lYThrSerAlaAsnCysLysHisrTArgLys***GluMetSerLeu-----GlyThrLeuA 351
DB 884 GAGAGCGCTCAC-----CAACATCTCAGCGGTGAATCAGACTTCAGACAGAGAGTACTTC 937
QY 351 spMetProGlyProHisGln 357
DB 938 AGCAA-----CACGAG 948

```

RESULT 10
 AAH16158
 ID AAH16158 standard; cDNA; 2224 BP.

XX AAH16158;
 XX AC
 XX 26-JUN-2001 (first entry)
 DT Human cDNA sequence SEQ ID NO:14925.
 DE
 XX

```

KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EF1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000BP-00116126.
XX
XX 29-JUL-1999; 990P-00248036.
XX 29-AUG-1999; 990P-00300253.
XX 11-JAN-2000; 2000BP-00118776.
XX 02-MAY-2000; 2000BP-00183767.
XX 09-JUN-2000; 2000BP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR MPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
PS
XX Claim 8; SEQ ID NO 14925; 2537bp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primer sets are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2224 BP; 419 A; 656 C; 806 G; 343 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	Length:	2224
Score:	631.50	142
Percent Similarity:	69.12%	8
Best Local Similarity:	65.44%	34
Query Match:	34.58%	33
DB:	4	3

FIGURE9 (1-361) x AAH16158 (1-2224)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPhe 116
DB 284 TTTCGATCGTGGCGAAGTCCGGCAGAGGCGACGGGCAAGAGTACGAGCCAACTTC 343
QY 117 IleLysLysArgAGLeuProSerSerArgArgGlyAlaSerArgGluGluIleGluArg 136
DB 344 ATCAAGAGAGCGCCCGCTGATCCAGCGCGGTGGGTGAGCCGGGAGAGATCGAGCGG 403

```

```

QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 404 GAGGTGAACATCTCTGGGGAGATCCGGACCCCAACATCATCACTCTGACGACGATCTTC 463
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 464 GAGAACAAAGACGACGCTGCTCATCTCGAGCTGATCTGCTGGGAGGCTCTTTCAC 523
QY 177 PheLeuAlaGluLysAspHis***-GlnArgMetArgProArgSerSerSerArgSe 196
DB 524 TTCTTGGCGGAGAAAGAGTGGCTGACGAGAGAGAGGCCACAGTTCTCTCAAGCAGATC 583
QY 196 rTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 584 CTGACACGGCTGACACTGCTGACCTTAAGCGCATCGACACCTTACCTGAAGCCGGA 643
QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 644 AACATCATGCTGCTGGACAAAGACGTCCCAACCCAGATCAAGCTCATCGACTTCGGC 703
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 704 ATCGCGCAAGATCGACGGCGGAGACGAGTTCAAGAACCA-----TC 745
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
DB 746 TTCGGACACCCCGAGGTTGTGGC----- 768
QY 276 euThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
DB 769 -----CCAGAGATTGGAACTATGAGCCGCTG 796
QY 296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisPro 311
DB 797 GGCTTGAGGCGGACATGTGGACATCGGTGCTACCTATATCTCT 843

```

RESULT 11
AAH78068
ID AAH78068 standard; DNA; 2224 BP.
AC AAH78068;
DT 26-NOV-2001 (first entry)
DE Nucleotide sequence of a human protein kinase/protein phosphatase.
KW Human; protein kinase; protein phosphatase; signal transduction;
KW intracellular signalling pathway; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 215..1579
FT /tag= a
FT /product= "protein kinase/protein phosphatase"
XX
XX MO200109345-A1.
XX
XX 08-FEB-2001.
XX
XX 28-JUL-2000; 2000WO-JP005060.
XX
XX 29-JUL-1999; 99JP-00248036.
XX 18-OCT-1999; 99US-0159590P.
XX 11-JAN-2000; 2000JP-00118776.
XX 17-FEB-2000; 2000US-0183322P.
XX 02-MAY-2000; 2000JP-00183767.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S,
XX Senoo C, Nezu J;

```

XX WP1: 2001-564736/63.
DR P-PSDB; AAG67425.
XX
PT New genes encoding protein kinase and protein phosphatase, useful for
PT identifying modulators which can be used to treat human or animal
PT disorders associated with the expression or function of these enzymes.
XX
XX Claim 1; Page 119-125; 336pp; Japanese.
XX
CC The present sequence encodes a human protein kinase/protein phosphatase.
CC The polypeptides are expected to participate in signal transduction in
CC cells. The kinase phosphatases are connected with intracellular
CC signaling pathways. Antisense oligonucleotides and compounds identified
CC by screening (agonists or antagonists) can be used to treat human or
CC animal disorders associated with the expression or function of the
CC protein. In addition, the polypeptides may be used as target molecules
CC for drug development
XX
SQ Sequence 2224 BP; 419 A; 656 C; 806 G; 343 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	8,99e-29	2224	142	8	33	3
Percent Similarity:	631.50					
Best Local Similarity:	69.12%					
Query Match:	65.44%					
DB:	34.58%					

FIGURE9 (1-361) x AAH78068 (1-2224)

```

QY 99 PheAlaIleValArgLysCys-----LysGlyThrGlyMetGlyTyrAlaAlaLysPhe 116
DB 284 TTTCGACATCGCGGAGATGCGCGAGAAAGGCGACGCGGACAGAGTACGCGCCAGCTTC 343
QY 117 IleLysLysArgGlnLeuProSerSerArgGlyValaSerArgGluIleGlyArg 136
DB 344 ATCAAGAAAGCGCGCTTCATCCAGCGCGGTGGGAGGACCGGAGAGATCGACCGG 403
QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
DB 404 GAGGTGAACATCTCTGGGGAGATCCGGACCCCAACATCATCACTCTGACGACGATCTTC 463
QY 157 GluAsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAsp 176
DB 464 GAGAACAAAGACGACGCTGCTCATCTCGAGCTGATCTGCTGGGAGGCTCTTTCAC 523
QY 177 PheLeuAlaGluLysAspHis***-GlnArgMetArgProArgSerSerSerArgSe 196
DB 524 TTCTTGGCGGAGAAAGAGTGGCTGACGAGAGAGGCCACAGTTCTCTCAAGCAGATC 583
QY 196 rTPThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArgAr 216
DB 584 CTGACACGGCTGACACTGCTGACCTTAAGCGCATCGACACCTTACCTGAAGCCGGA 643
QY 216 gThrSerCysCysTrpThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
DB 644 AACATCATGCTGCTGGACAAAGACGTCCCAACCCAGATCAAGCTCATCGACTTCGGC 703
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
DB 704 ATCGCGCAAGATCGACGGCGGAGACGAGTTCAAGAACCA-----TC 745
QY 256 rSerVal-ArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGlyL 276
DB 746 TTCGGACACCCCGAGGTTGTGGC----- 768
QY 276 euThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
DB 769 -----CCAGAGATTGGAACTATGAGCCGCTG 796
QY 296 rPLeuGlySer**TyrValGluHisArgArgHisHisLeuHisPro 311

```

Db 797 GGCGTGGAGGGGACATGTGGAGCATGGTGTATCATCACTATATCTT 843

RESULT 12

AAK93262

ID AAK93262 standard; cDNA; 757 BP.

XX AAK93262;

XX 06-NOV-2001 (first entry)

XX Human cDNA clone representative sequence, SEQ ID NO: 1722.

XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

XX in genetic manipulation.

XX Example 11; SEQ ID NO 1722; 1380BP + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA

XX clones. 830 cDNA molecules encoding a human protein have been isolated

XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have

XX been determined. Primers for synthesizing the full length cDNA are useful

XX for clarifying the function of the protein encoded by the cDNA. The full

XX length clones were obtained by construction of full length enriched cDNA

XX libraries that were synthesized by the oligo-capping method. The primers

XX enable the production of the full length cDNA easily without any special

XX methods. The present sequence was used as the representative sequence

XX from a human clone which was used in homology searches to identify the

XX clone. Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in CD-ROM format directly from

XX EPO

XX Sequence 757 BP; 176 A; 219 C; 220 G; 133 T; 0 U; 9 Other;

XX

Alignment Scores:

Pred. No.: 2 67e-26 Length: 757

Score: 581.50 Matches: 135

Percent Similarity: 66.37% Conservative: 15

Best Local Similarity: 59.73% Mismatches: 41

Query Match: 31.85% Indels: 35

DB: 4 Gaps: 3

FIGURE 9 (1-361) x AAK93262 (1-757)

QY 99 PheAlAlIleValAlArgIlyscys-----IysGIYThGlyMetGluTyrAlaAlaIysPhe 116

Db 139 TTTCGCAATCGTGGGAGATGCCCGCAGAAAGGCGACGCGAAGAGTACCCACCAAGTTTC 198

QY 117 IleLysIlyArGrhneuproserserArGrGlyValSerArGrGluGluIleGluArg 136

Db 199 ATCAAGAGAGCGCGCTGTTCATCATCCAGCGCGGTGAGCCGGAGAGATCGAGCGG 258

QY 137 GluValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPhe 156

Db 259 GAGGTGAACATCTCGGAGATCCGACCCCAACATCATCATCCCTCACACATCTTC 318

QY 157 GluAsnIlyThrAspValValIleuIleLeuGluLeuValSerGIlyGIlyLeuPheAsp 176

Db 319 GAGAACAGACGACGAGCTGTCTCTCATCTCGAGCTGTCTCTGGGGAGAGCTCTTTGAC 378

QY 177 PheLeuAlaGluIlyAspHis***GlnArgYerArGrProArGrSerSerSerArgSe 196

Db 379 TTCCCTGGCGGAGAGAGTCCGTGACGAGGACGAGCCACCCAGTTCTCTCAGCAGATC 438

QY 196 rTrpThrValSerThrThrCysThrProSerAlaSerArGrThrLeuThr***SerArGr 216

Db 439 CTGGACGGCGTTCATCTTACCTGCACTTAAAGGCATGCGACACTTGAACCTGAGCGGAA 498

QY 216 gThrSerCysCysTrpThrIserMetGlnProAlaHisAlaLeuSerSerThrLeuAl 236

Db 499 AACATCATGCTGCTGACACAGACGTGCCCAACCCAGCATCAAGCTCATGCACTTCGGC 558

QY 236 aSerArGrThrGrIySerArGrProValAlaSerSerArGrThrSerLeuAlaArgGlnSerSe 256

Db 559 ATCGCNCACAGATGANGCGGGGAGACGAATTCANAAACATCTTCGGACCCCGAATTT 618

QY 256 rSerValArgGlyArgCysGIyHisIleProIleGIy***IleLeuHisGIlyLeuGIlyLe 276

Db 619 TGTGGCCCCAANAATTGTGAACCTATGAAACCGCTGGC----- 655

QY 276 uThrCysLeuAsnAsnProValPheHisSerPro***AspCysIlyLeu***ThrThrTr 296

Db 655 ----- 655

QY 296 PleuGIySer***TyrValGIyHis-ArgArgHisIleLeuHisProValSerAla***A 316

Db 656 -CTTGAGCNGCAGCATGCTGAACATCTCT-----GTCTCCCTCATATCTCTCTGAA 705

QY 316 spGIyGlnGIlyPro 320

Db 706 CCGGTGCATCCCCC 719

RESULT 13

AAK91856

ID AAK91856 standard; cDNA; 757 BP.

XX AAK91856;

XX 06-NOV-2001 (first entry)

XX Human cDNA 5'-end sequence, SEQ ID NO: 316.

XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

XX in genetic manipulation.

PS Claim 2; SEQ ID NO 316; 1380bp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is the nucleotide sequence of the 5'-end of
 CC a cDNA provided in the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in CD-
 CC ROM format directly from EPO

XX
 SQ Sequence 757 BP; 176 A; 219 C; 220 G; 133 T; 0 U; 9 Other;

Alignment Scores:
 Pred. No.: 2 67e-26 Length: 757
 Score: 581.50 Matches: 135
 Percent Similarity: 66.37% Conservative: 15
 Best Local Similarity: 59.73% Mismatches: 41
 Query Match: 31.85% Indels: 35
 DB: 4 Gaps: 3

FIGURE 9 (1-361) x AAK91856 (1-757)

QY 99 PhenAlaIleValArgLysCys-----LysGlyThrGlyMetGlnTyrAlaAlaLysPhe 116
 Db 139 TTTCGATCGTCGGAGATCGCGCAAGAGGACCGGCAAGAGTACGACGACGATTC 198
 QY 117 IleLysValArgLysPheProSerSerArgArgGlyValSerArgGlnGlnIleGlnArg 136
 Db 199 ATCAAGAGCGCGCGCTGTCATCAGCGCGCGCTGGGTAGCGCGGAGAGATCGAGCGG 258
 QY 137 GluValSerIleLeuArgGlnIleArgHisProAsnIleIleThrLeuHisAspValPhe 156
 Db 259 GAGGTGAACATCCCTCGCGAGATCCGCGACCCCAACATCACTCCCTGCGACGACATCTTC 318
 QY 157 GluAsnLysThrAspValIleLeuIleLeuGlnLeuValSerGlyGlnLeuPheAsp 176
 Db 319 GAGAACAGAGCGAGCGTGTCTCATCTGAGTGTCTCTGCGGAGGCTTTGAC 378
 QY 177 PheLeuAlaGlnLysAspHis***GlnArgMetArgProArgSerSerSerSerArg 196
 Db 379 TTCTGCGGAGAGAGATCGCTGCGTGCAGAGACGACCCAGTTCTCTCAAGCAGATC 438
 QY 196 rTrrThrValSerThrThrCysThrProSerAlaSerArgThrLeuThr***SerArg 216
 Db 439 CTGACGCGCGCTTCACTCTGCTGACTTGAAGCGCATCGCACCTTTGACTGAAGCGGAA 498
 QY 216 gThrSerCysSerThrProSerMetGlnProAlaHisAlaLeuSerSerSerThrLeu 236
 Db 499 AACATCATGCTGCTGCGACAGAGAGTGTCCCAACCCAGAACAGTCACTGCTCGG 558
 QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSer 256
 Db 559 ATCGACNCAAGATCGAGGCGGAGCAATTCANAACTCTCGCACCCCGGAATT 618
 QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly***IleLeuHisGlyLeuGly 276
 Db 619 TGTGGCCCCAANAATTGTGAACATGTGAACCGCTGGC----- 655
 QY 276 uThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrTyr 296
 Db 655 ----- 655
 QY 296 pLeuGlySer***TyrValGlnHis-ArgArgHisHisLeuHisProValSerAla***A 316
 Db 656 -CTTGAAGCGAGCATGTGAACATCCGT-----GTCCTCCCTAATCTCTCGAA 705
 QY 316 spGlyGlnGlyPro 320
 Db -----

Db 706 CCGGTGCATCCCC 719
 RESULT 14
 ID AAK70641 standard; DNA; 12638 BP.
 XX AAK70641;
 AC
 DT 06-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25453.
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW Cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227183P.
 PR 22-AUG-2000; 2000US-0227009P.
 PR 23-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.

```
PR 08-SEP-2000; 2000US-0232081P
PR 12-SEP-2000; 2000US-0231968P
PR 14-SEP-2000; 2000US-0232357P
PR 14-SEP-2000; 2000US-0232398P
PR 14-SEP-2000; 2000US-0232399P
PR 14-SEP-2000; 2000US-0232400P
PR 14-SEP-2000; 2000US-0232401P
PR 14-SEP-2000; 2000US-0233063P
PR 14-SEP-2000; 2000US-0233064P
PR 14-SEP-2000; 2000US-0233065P
PR 21-SEP-2000; 2000US-0234223P
PR 21-SEP-2000; 2000US-0234274P
PR 25-SEP-2000; 2000US-0234997P
PR 25-SEP-2000; 2000US-0234998P
PR 25-SEP-2000; 2000US-0235484P
PR 27-SEP-2000; 2000US-0235834P
PR 27-SEP-2000; 2000US-0235836P
PR 29-SEP-2000; 2000US-0236377P
PR 29-SEP-2000; 2000US-0236378P
PR 29-SEP-2000; 2000US-0236386P
PR 29-SEP-2000; 2000US-0236389P
PR 29-SEP-2000; 2000US-0236397P
PR 02-OCT-2000; 2000US-0236802P
PR 02-OCT-2000; 2000US-0237037P
PR 02-OCT-2000; 2000US-0237038P
PR 02-OCT-2000; 2000US-0237039P
PR 02-OCT-2000; 2000US-0237040P
PR 13-OCT-2000; 2000US-0239935P
PR 13-OCT-2000; 2000US-0239937P
PR 20-OCT-2000; 2000US-0240960P
PR 20-OCT-2000; 2000US-0241221P
PR 20-OCT-2000; 2000US-0241785P
PR 20-OCT-2000; 2000US-0241786P
PR 20-OCT-2000; 2000US-0241787P
PR 20-OCT-2000; 2000US-0241808P
PR 20-OCT-2000; 2000US-0241809P
PR 20-OCT-2000; 2000US-0241826P
PR 01-NOV-2000; 2000US-0244617P
PR 08-NOV-2000; 2000US-0246474P
PR 08-NOV-2000; 2000US-0246475P
PR 08-NOV-2000; 2000US-0246476P
PR 08-NOV-2000; 2000US-0246477P
PR 08-NOV-2000; 2000US-0246478P
PR 08-NOV-2000; 2000US-0246532P
PR 08-NOV-2000; 2000US-0246534P
PR 08-NOV-2000; 2000US-0246534P
PR 08-NOV-2000; 2000US-0246535P
PR 08-NOV-2000; 2000US-0246536P
PR 08-NOV-2000; 2000US-0246537P
PR 08-NOV-2000; 2000US-0246538P
PR 08-NOV-2000; 2000US-0246539P
PR 08-NOV-2000; 2000US-0246540P
PR 08-NOV-2000; 2000US-0246610P
PR 08-NOV-2000; 2000US-0246611P
PR 08-NOV-2000; 2000US-0246613P
PR 17-NOV-2000; 2000US-0249207P
PR 17-NOV-2000; 2000US-0249208P
PR 17-NOV-2000; 2000US-0249209P
PR 17-NOV-2000; 2000US-0249210P
PR 17-NOV-2000; 2000US-0249211P
PR 17-NOV-2000; 2000US-0249212P
PR 17-NOV-2000; 2000US-0249213P
PR 17-NOV-2000; 2000US-0249214P
PR 17-NOV-2000; 2000US-0249215P
PR 17-NOV-2000; 2000US-0249216P
PR 17-NOV-2000; 2000US-0249217P
PR 17-NOV-2000; 2000US-0249218P
PR 17-NOV-2000; 2000US-0249244P
PR 17-NOV-2000; 2000US-0249245P
PR 17-NOV-2000; 2000US-0249246P
PR 17-NOV-2000; 2000US-0249255P
PR 17-NOV-2000; 2000US-0249257P
PR 17-NOV-2000; 2000US-0249259P
PR 17-NOV-2000; 2000US-0249300P

PR 01-DEC-2000; 2000US-0250160P
PR 01-DEC-2000; 2000US-0250391P
PR 05-DEC-2000; 2000US-0251030P
PR 05-DEC-2000; 2000US-0251988P
PR 05-DEC-2000; 2000US-0256719P
PR 06-DEC-2000; 2000US-0251479P
PR 06-DEC-2000; 2000US-0251856P
PR 08-DEC-2000; 2000US-0251868P
PR 08-DEC-2000; 2000US-0251869P
PR 08-DEC-2000; 2000US-0251989P
PR 11-DEC-2000; 2000US-0254097P
PR 05-JAN-2001; 2001US-0259678P

PA (HUMA-) HUMAN GENOME SCT INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
DR
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PS Disclosure; SEQ ID NO 25453; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX
SQ Sequence 12638 BP; 2397 A; 3383 C; 3994 G; 2864 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,766-23 Length: 12638
Score: 548.00 Matches: 152
Percent Similarity: 47.81% Conservative: 12
Best Local Similarity: 44.31% Mismatches: 61
Query Match: 30.01% Indels: 119
DB: 4 Gaps: 9

FIGURE9 (1-361) x AAK70641 (1-12638)

QY 81 ValLeuPProSeSerGlySerAlaAlaLeuSerCysSerPheProHisSerGly----Phe 99
DB 6048 TTACTCCCTCCGCTGCTCAACCCCACTCACTCCCTCCCTCCAGGGCAAGTTT 6107
QY 100 AlaIleValArgLysCys-----LysGlyThrGlyMetGluTyrAlaAlaLysPheIle 117
DB 6108 GCGATCGTCGCGAAGTCCGCGCAGAAAGGACGCGCAAGAGTACGCCAAGTTCATC 6167
QY 118 LysLysAArgArgLeuProSerSerArgArgLysValSerArgGluGluIleGluArgGlu 137
DB 6168 AAGAAGCCCGCGCTGTCTCATCCAGCGCGGTGGGTGAGCCCGGAGGATCGACCGGGAG 6227
QY 138 ValSerIleLeuArgGluIleArgHisProAsnIleIleThrLeuHisAspValPheGlu 157
DB 6228 GTGAACATCTCGCGGAGATCCGCGCAACCATCATCACCCTCGACAGCATCTTCAG 6287
QY 158 AsnLysThrAspValValLeuIleLeuGluLeuValSerGlyGlyGluLeuPheAspPhe 177
```



```
QY 196 rTTPThrValSerThrCysThrProSerAlaSerArgThrLeuThr**SerArgAr 216
Db 540 CTGGATGGGGTGAACCTACCTCACAAGAAAATTGCTCACTTGTATCTCAAGCCGANA 599
QY 216 gThrSerCysCysTTPThrSerMetGlnProAlaHisAlaLeuSerSerSerThrLeuAl 236
Db 600 AACATTATGTTGTAGACAGAAATATTCCTCCATCCACATCAAGCTGATGACTTTGGT 659
QY 236 aSerArgThrGlySerArgProValAlaSerSerArgThrSerLeuAlaArgGlnSerSe 256
Db 660 CTGGCTCAGCAAAATAGAAAGATGAGATTGAATTAAAG----- 696
QY 256 rSerValArgGlyArgCysGlyHisHisProIleGly**IleLeuHisGly-LeuGlyL 276
Db 697 -----ATATTTTGGAGCCGANA 716
QY 276 euThrCysLeuAsnAsnProValPheHisSerPro**AspCysLysLeu**ThrThrT 296
Db 717 TTGTGTGCTCCAGAAAT-----TGTGAATACGAGCCCTGT 752
QY 296 rPleuGlySer**TyValGlnHisArgArgHisHisLeuHisProVal 312
Db 753 GGTCTGAGGCTGACATGTGAGCATAGGGGTCATCACCCTACCTCCTT 802
```

Search completed: July 1, 2004, 01:44:14
Job time : 494 secs

